

LB 164

CRFE

Access DB# 79000

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: M.A. WALICKA Examiner #: 78201 Date: Oct. 29 '02
 Art Unit: 1652 Phone Number 305-7270 Serial Number: 09/724,571
 Mail Box and Bldg/Room Location: 10D06 Results Format Preferred (circle): PAPER DISK E-MAIL
10D01

If more than one search is submitted, please prioritize searches in order of need. me

Please provide a detailed statement of the search topic; and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Beta-secretase enzyme composition and method
 Inventors (please provide full names): Anderson

Earliest Priority Filing Date: 06/15/1999

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search:

SEQ ID NO: 75

103 } these are polypeptides
 104 } of 4 AA
 72 } these are polypeptides
 78 } of 8-15 AA
 81 }
 91 }
 103-361
 104-4
 72-14
 78-7
 81-9
 91-9

Thank you in advance.

Walicka

RECEIVED
 OCT 30 2002
 Edward Hant
 Technical Info. Specialist
 STIC/Biotech
 CMI 6B02 Tel: 305-9203

STAFF USE ONLY

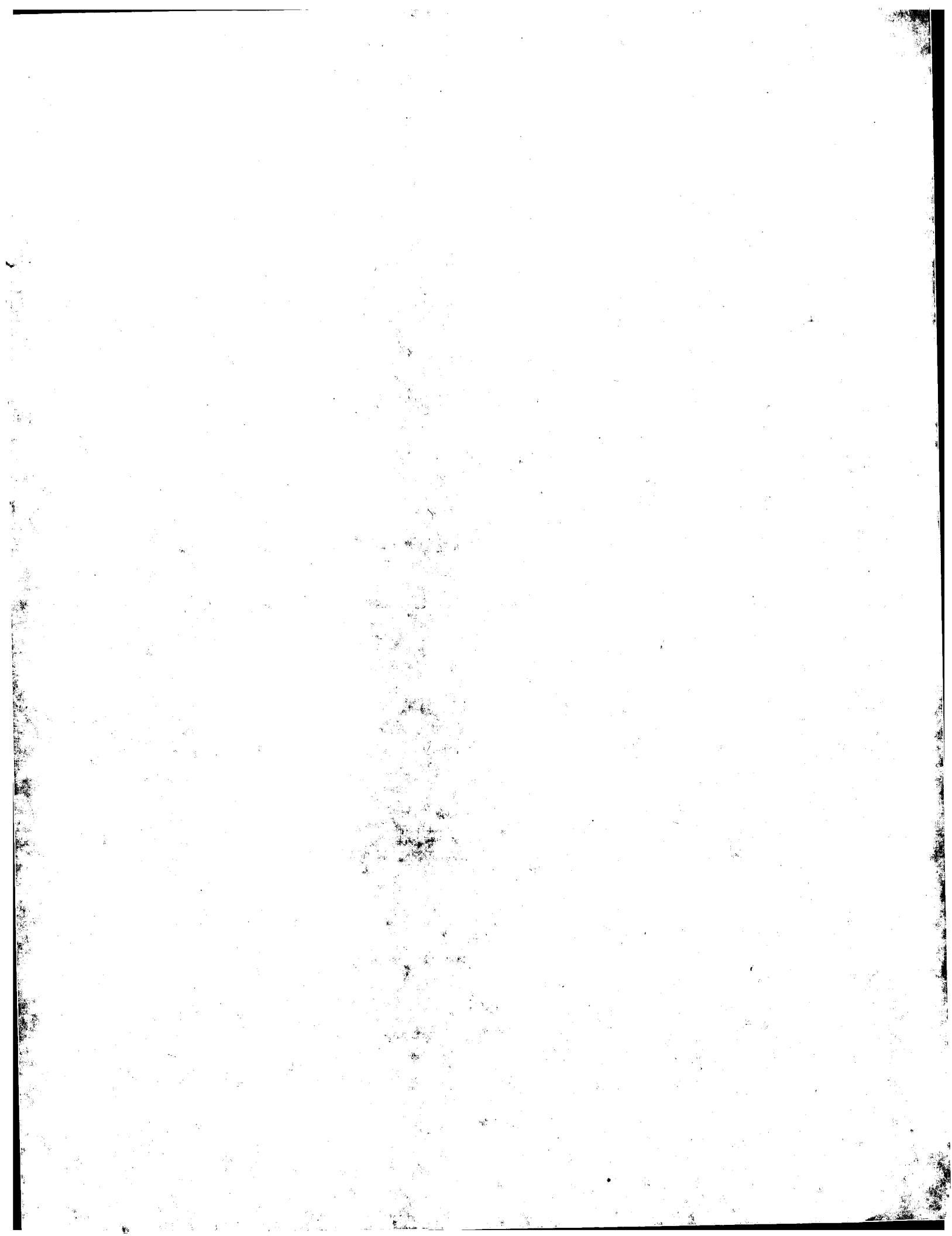
Searcher: _____
 Searcher Phone #: _____
 Searcher Location: _____
 Date Searcher Picked Up: 10/30/02
 Date Completed: 10/30/02
 Searcher Prep & Review Time: _____
 Clerical Prep Time: _____
 Online Time: _____

Type of Search

NA Sequence (#) _____
 AA Sequence (#) 7
 Structure (#) _____
 Bibliographic _____
 Litigation _____
 Fulltext _____
 Patent Family _____
 Other _____

Vendors and cost where applicable

STN _____
 Dialog _____
 Questel/Orbit _____
 Dr.Link _____
 Lexis/Nexis _____
 Sequence Systems 02
 WWW/Internet _____
 Other (specify) _____



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:20:50 ; Search time 1.60197 Seconds
(without alignments)
277.344 Million cell updates/sec

Title: US-09-724-571-104
Perfect score: 20
Sequence: 1 VNED 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	4	21	AA1980.DAT
2	20	100.0	5	18	AA1981.DAT
3	20	100.0	5	19	AA1982.DAT
4	20	100.0	5	20	AA1983.DAT
5	20	100.0	5	22	AA1984.DAT
6	20	100.0	5	22	AA1985.DAT
7	20	100.0	5	22	AA1986.DAT
8	20	100.0	5	22	AA1987.DAT
9	20	100.0	5	22	AA1988.DAT
10	20	100.0	5	22	AA1989.DAT
11	20	100.0	5	22	AA1990.DAT

12	20	100.0	9	21	AA1991.DAT	A peptide fragment
13	20	100.0	9	21	AA1992.DAT	Substrate for beta
14	20	100.0	9	22	AA1993.DAT	Protease inhibitor
15	20	100.0	10	18	AA1994.DAT	Beta-secretase sub
16	20	100.0	10	20	AA1995.DAT	Synthetic oligopep
17	20	100.0	10	21	AA1996.DAT	Beta-APP alpha-sec
18	20	100.0	10	22	AA1997.DAT	Human APP-Sw beta
19	20	100.0	10	22	AA1998.DAT	Human amyloid prec
20	20	100.0	10	22	AA1999.DAT	Synthetic Asp2 rec
21	20	100.0	10	22	AA2000.DAT	Human beta-amyloid
22	20	100.0	10	22	AA2001.DAT	Human APP-Sw beta
23	20	100.0	10	22	AA2002.DAT	Peptide 5-5' SW, fo
24	20	100.0	10	22	AA2003.DAT	Synthetic peptide
25	20	100.0	10	22	AA2004.DAT	Swedish mutation p
26	20	100.0	11	22	AA2005.DAT	APP Swedish varian
27	20	100.0	11	22	AA2006.DAT	Asp 1 substrate se
28	20	100.0	11	22	AA2007.DAT	Asp2 substrate Swe
29	20	100.0	13	13	AA2008.DAT	Exon XII of human
30	20	100.0	15	15	AA2009.DAT	Hbeta-derived pept
31	20	100.0	15	22	AA2010.DAT	Human ribosome S7
32	20	100.0	16	21	AA2011.DAT	Human beta-amyloid
33	20	100.0	16	22	AA2012.DAT	Beta-amyloid precu
34	20	100.0	19	21	AA2013.DAT	Vada homologous pe
35	20	100.0	19	22	AA2014.DAT	Amyloid precursor
36	20	100.0	20	19	AA2015.DAT	Fluorogenic protea
37	20	100.0	20	20	AA2016.DAT	Pullanase type DBE
38	20	100.0	20	21	AA2017.DAT	Beta-APP alpha-sec
39	20	100.0	20	22	AA2018.DAT	Protease binding s
40	20	100.0	21	16	AA2019.DAT	Amylopectin starch
41	20	100.0	21	17	AA2020.DAT	Potato debranching
42	20	100.0	21	18	AA2021.DAT	Beta-secretase sub
43	20	100.0	21	19	AA2022.DAT	Fluorogenic protea
44	20	100.0	21	19	AA2023.DAT	Fluorogenic protea
45	20	100.0	21	20	AA2024.DAT	Synthetic oligopep

ALIGNMENTS

RESULT 1	AA1980.DAT
AA1981.DAT	standard; peptide: 4 AA.
AA1982.DAT	
AA1983.DAT	
AA1984.DAT	14-FEB-2001 (first entry)
AA1985.DAT	
AA1986.DAT	Amyloid-beta precursor protein Swedish mutation site.
AA1987.DAT	
AA1988.DAT	Aldehyde caspase inhibitor; antibody; amyloid-beta precursor;
AA1989.DAT	caspase cleavage; neurodegenerative disease.
AA1990.DAT	
AA1991.DAT	Synthetic.
AA1992.DAT	
AA1993.DAT	WO20063250-A1.
AA1994.DAT	26-OCT-2000.
AA1995.DAT	13-APR-2000; 2000WO-CA00414.
AA1996.DAT	15-APR-1999; 99US-0129495.
AA1997.DAT	(MERI) MERCK FROST CANADA & CO.
AA1998.DAT	Gervais F, Roy S, Nicholson DW, Xu D, Robertson G, Huang J;
AA1999.DAT	WPI, 2000-687160/67.
AA2000.DAT	Novel antibody that recognizes neo-epitopes of caspase cleaved
AA2001.DAT	amyloid-beta precursor protein useful for diagnosing conditions
AA2002.DAT	involving neuronal apoptosis such as Alzheimer's, Huntington's and
AA2003.DAT	Parkinson's disease -
AA2004.DAT	
AA2005.DAT	
AA2006.DAT	
AA2007.DAT	
AA2008.DAT	
AA2009.DAT	
AA2010.DAT	
AA2011.DAT	
AA2012.DAT	
AA2013.DAT	
AA2014.DAT	
AA2015.DAT	
AA2016.DAT	
AA2017.DAT	
AA2018.DAT	
AA2019.DAT	
AA2020.DAT	
AA2021.DAT	
AA2022.DAT	
AA2023.DAT	
AA2024.DAT	

PS Example 1; Page 7; 65pp; English.
XX
CC The present invention describes an antibody which recognises a
CC neo-epitope created by the caspase mediated cleavage of amyloid-beta
CC precursor protein (APP) or amyloid-beta precursor like protein (ABPP).
CC These proteins are involved in the premature death of cells at sites of
CC neurodegeneration. The antibody of the invention can thus be used in the
CC treatment of neurodegenerative diseases such as Alzheimer's, Huntington's
CC and Parkinson's diseases, amyotrophic lateral sclerosis, progressive
CC Creutzfeldt-Jacob disease, spongiform encephalopathy, Friedreich's ataxia,
CC fatal familial insomnia, Pelizaeus-Werzacher disease, schizophrenia,
CC dentatorubropallidoluyian atrophy, spinocerebellar atrophy type 3,
CC spinal bulbar muscular atrophy, spinal cord injury, stroke and brain
CC injury.
XX
SO Sequence 4 AA:

Query Match 100.0%; Score 20; DB 21; Length 4;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
Db 1 VNLD 4

RESULT 2
ID AAM08217 standard; peptide: 5 AA.
XX
AC AAM08217;
XX
DT 05-SEP-1997 (first entry)
XX
DE Swedish double mutant APP beta-cleavage site.
XX
KW Beta-cleavage site; beta amyloid precursor protein; APP; beta-secretase;
KW alpha-secretase; proteolytic cleavage; inhibitor; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO9640885-A2.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09985.
XX
PR 07-JUN-1995; 95US-0485152.
PR 07-JUN-1995; 95US-0480498.
XX
PA (ATHE-) ATHENA NEUROSCIENCES INC.
XX
PI Anderson JP, Chrysler SMS, Jacobson-croak KL, Kelm PS;
PI Mcconlogue LC, Sinha S, Tan H;
XX
PT WPI: 1997-052304/05.
XX
PT Beta-secretase which specifically cleaves beta-amyloid precursor
PT protein - useful to screen for inhibitors useful in treatment of
PT Alzheimer's disease
XX
PS Claim 5; Page 60; 92pp; English.
XX
CC AAM08216, AAM08217 and AAM08350 represent beta-cleavage sites from
CC beta-amyloid precursor proteins (APP). These sequences are recognised by
CC the enzyme of the invention. The enzyme of the invention is
CC beta-secretase, and specifically cleaves beta-APP at one of these sites.
CC Normal processing of beta-APP is thought to occur via cleavage between
CC residues 16 and 17 of the beta-amyloid peptide region by an
CC alpha-secretase. Pathogenic processing is thought to occur by
CC beta-secretase cleavage of beta-APP. Beta-secretase activity can be
CC detected and measured using a method of the invention, which detects at

CC least one of the beta-secretase cleavage products formed on cleavage. The
CC method can be used to determine whether a test substance inhibits
CC proteolytic cleavage, by beta-secretase, of beta-APP. Compounds effective
CC to at least partially inhibit beta-secretase activity can be used to
CC inhibit cleavage of beta-APP in cells or mammalian hosts. Isolation and
CC purification of beta-secretase will permit chemical modelling of a
CC critical event in the pathology of Alzheimer's disease.
XX
SO Sequence 5 AA:

Query Match 100.0%; Score 20; DB 18; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
Db 1 VNLD 4

RESULT 3
ID AAM61151 standard; Peptide: 5 AA.
XX
AC AAM61151;
XX
DT 26-OCT-1998 (first entry)
XX
DE APP Swedish double mutation cleavage site.
XX
KW Beta-secretase; human; beta-amyloid precursor protein; APP;
KW protease; inhibitor; screening; Alzheimer's disease; therapy.
XX
OS Homo sapiens.
XX
PN WO9826059-A1.
XX
PD 18-JUN-1998.
XX
PF 11-DEC-1996; 96WO-US19549.
XX
PR 11-DEC-1996; 96WO-US19549.
XX
PA (ATHE-) ATHENA NEUROSCIENCES INC.
XX
PI Anderson JP, Chrysler SMS, Kelm PS, Sinha S;
XX
PT WPI: 1998-348519/30.
XX
PT Novel beta-secretase which cleaves beta-amyloid precursor protein -
PT are useful for treating for compounds which inhibit the cleavage and
PT are useful for treating Alzheimer's disease
XX
PS Disclosure: Page 20; 39pp; English.
XX
CC This peptide comprises the site of the 'Swedish' double mutation
CC beta-amyloid precursor protein (APP) (MPP-C125 SW) that is cleaved
CC by a novel beta-secretase isolated from human 293 cells. This
CC protease cleaves APP at the N-terminus of the beta-amyloid peptide
CC (beta-AP) and is believed to be the putative beta-secretase
CC responsible for the pathogenic processing of APP to beta-AP in
CC Alzheimer's disease, Down's syndrome and HCHWA-D. Recombinant
CC 125 amino acids of APP (wild-type (see AAM61150) or Swedish double
CC mutation) fused to the C-terminal end of maltose binding protein.
CC The fusion proteins were expressed in Escherichia coli, and used as
CC substrates for beta-secretase in beta-secretase inhibitor assays.
CC Compounds that inhibit APP cleavage by beta-secretase may be useful
CC in the treatment of Alzheimer's disease.
XX
SO Sequence 5 AA:

Query Match 100.0%; Score 20; DB 19; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;

Matches 4: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNLD 4
|||||
Db 1 VNLD 4

RESULT 4
AAV33751
ID AAV33751 standard; Protein; 5 AA.

AC AAY33751;

DT 09-NOV-1999 (first entry)

DE Swedish mutant beta-amyloid protein precursor (APP) cleavage site.

KW Beta-secretase; beta-amyloid protein precursor; APP; Down's syndrome;
KM Alzheimer's disease; cleavage site; mutant.

OS Homo sapiens.
OS Synthetic.

PN US5942400-A.

PD 24-AUG-1999.

PF 07-JUN-1996; 9605-0659984.

PR 07-JUN-1996; 9605-0659984.

PR 07-JUN-1995; 9505-0480498.

PR 07-JUN-1995; 9505-0485152.

PA (ELAN-) ELAN PHARM INC.

PI Anderson JP, Jacobson-Croak KL, Sinha S;

DR WPI; 1999-517417/43.

PT A method for detecting human beta-secretase cleavage of polypeptides
PT useful for identifying beta-secretase inhibitors

PS Examples; Column 28; 43pp; English.

CC This sequence is the Swedish mutant beta-amyloid protein precursor (APP)
CC cleavage site. APP is cleaved by beta-secretase AAY33741. The wild type
CC cleavage site AAY33750 and the Swedish mutant version are used in a
CC method for detecting human beta-secretase cleavage of polypeptides and
CC for identifying beta-secretase inhibitors. Inhibition of beta-secretase
CC activity would be useful for chemical modelling of a critical event in
CC the pathology of Alzheimer's disease. Inhibitors of beta-secretase would
CC be useful for the prevention and treatment of Alzheimer's disease and
CC Down's Syndrome.

SO Sequence 5 AA;

Query Match 100.0%; Score 20; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNLD 4
|||||
Db 1 VNLD 4

RESULT 5
AAB47261

ID AAB47261 standard; Peptide; 5 AA.

AC AAB47261;

DT 18-JUL-2001 (first entry)

DE Swedish mutation APP sequence for cleavage by beta-secretase.

KW Beta-secretase; isotype; beta-amyloid precursor protein; APP;
KM beta-amyloid peptide; beta-AP; Alzheimer's disease; Downs syndrome;
KM HCHWA-D; Swedish mutation; maltose binding protein; MBP.

OS Homo sapiens.

PN US6221645-B1.

PD 24-APR-2001.

PF 07-JUN-1996; 9605-0660531.

PR 07-JUN-1995; 9505-0480498.

PA (ELAN-) ELAN PHARM INC.

PI Chrysler SMS, Sinha S, Keim PS, Anderson JP, Tan H, McConlogue LC;

DR WPI; 2001-315578/33.

PT Novel antibody that specifically binds native beta-secretase protein,
PT useful for raising anti-idiotypic antibodies and for detecting or
PT diagnosing pathological conditions related to presence of respective
PT antigens

PS Example; Column 28; 42pp; English.

CC The sequences given in AAB47260-61 represent cleavage sites derived
CC from wild-type and the Swedish mutation of beta-amyloid precursor
CC protein (APP). These cleavage sites were used in fusion proteins
CC which were used as substrates for the beta-secretase protein which
CC is characterized by an ability to cleave the 695-amino acid isotype
CC of APP between amino acids 596 and 597. The fusion proteins contain
CC the carboxy-terminal end of Maltose binding protein (MBP) fused to
CC the carboxy-terminal 125 amino acids of either wild type APP or APP
CC containing the Swedish mutation. Beta-secretase is thought to be
CC responsible for the pathogenic processing of APP to form beta amyloid
CC peptide (beta-AP) in beta-AP related conditions, e.g. Alzheimer's
CC disease, Downs syndrome, HCHWA-D etc. Beta-secretase has a molecular
CC weight of 260-300 kD and will bind to wheat germ agglutinin but not to
CC concanavalin A. Beta-secretase will cleave both the wild type and
CC the Swedish mutation of APP.

SO Sequence 5 AA;

Query Match 100.0%; Score 20; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNLD 4
|||||
Db 1 VNLD 4

RESULT 6
AAY94771

ID AAY94771 standard; Protein; 8 AA.

AC AAY94771;

DT 12-FEB-2001 (first entry)

DE Beta-secretase substrate peptide SEQ ID 17.

KW Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease;

KM Down's syndrome; amyloid angiopathy; gene therapy; neuroprotective.

OS Synthetic.

PN WO200058479-A1.

```

PD 05-OCT-2000.
PF 23-MAR-2000; 2000WO-US07755.
PR 26-MAR-1999; 99US-0277229.
PA (AMGE-) AMGEN INC.
PI Citron M, Vassar RJ, Bennett BD;
XX WPI: 2000-594643/56.
XX Isolated beta-secretase nucleic acids and encoded polypeptides, useful
PT for diagnosis and gene therapy of Alzheimer's disease
XX Example 10; Page 117; 145pp; English.
XX This invention relates to 3 nucleotide sequences encoding beta-secretase
CC proteins. Beta-secretase is an enzyme involved in the production of one
CC of the components of amyloid plaques involved in Alzheimer's disease. The
CC invention includes an expression vector comprising the nucleotide
CC sequence, a host cell comprising the expression vector, and a process for
CC producing the protein through culturing the transformed cells. Also
CC included in the invention are a polypeptide derivative of the
CC beta-secretase protein, a fusion protein comprising beta-secretase fused
CC to a heterologous amino acid sequence, and a method for modulating the
CC levels of beta-secretase polypeptide in a mammal comprising administering
CC the polynucleotide sequence. Beta-secretase exhibits neuroprotective and
CC neurotrophic activity. The beta-secretase gene and related genes on chromosomes
CC map locations of the beta-secretase gene and related genes on chromosomes
CC as hybridization probes in diagnostic assays to test for the presence
CC of beta-secretase DNA or RNA, such as in Alzheimer's disease. Down's
CC syndrome, and amyloid angiopathy. The nucleotide sequence may also be
CC used as anti-sense inhibitors of beta-secretase expression, in gene
CC therapy of Alzheimer's disease, and for the identification of compounds
CC that modulate beta-secretase activity. Antibodies to the beta-secretase
CC protein may be used for in vitro and in vivo diagnostic purposes to
CC detect the presence of beta-secretase polypeptide in a body fluid or cell
CC sample. The present sequence represents a beta-secretase substrate
CC peptide.
XX
SQ Sequence 8 AA;
QY 1 VNLD 4      100.0%; Score 20; DB 21; Length 8;
   ||||         Pred. No. 6.4e+05;
Db 2 VNLD 5      Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps .0.

RESULT 7
ID AA068473 standard; Peptide; 8 AA.
XX AA068473;
DT 16-JAN-2002 (first entry)
XX Human Breast cancer-associated protein isoform, BPI-268 peptide #2.
DE Human Breast cancer-associated protein isoform; breast cancer;
XX Human; Breast cancer-associated protein isoform; breast cancer;
KW immunogen; cytostatic; BPI; tryptic digest peptide.
OS Homo sapiens.
XX WO200171357-A2.
PN 27-SEP-2001.
XX 20-MAR-2001; 2001WO-GB01219.
PF
XX

```

PR	20-MAR-2000;	2000GB-0006695.	
XX	24-MAR-2000;	2000GB-0007265.	
XX			
PA	(OXFO-) OXFORD GLYCOSCIENCES UK LTD.		
XX			
PI	Herath HMC, O'Hare MJ, Page MJ, Parekh RB, Waterfield MD;		
XX			
DR	WPI: 2001-611532/70.		
PT	Identifying proteins for clinical screening, diagnosis and prognosis of		
PT	breast cancer, comprises detecting Breast Cancer-Associated Protein		
PT	Isoforms (Bpis) using two-dimensional electrophoresis -		
XX			
PS	Claim 9; Page 53; 197pp; English.		
XX			
CC	The invention relates to diagnosing, determining the stage or severity,		
CC	or identifying the risk of a subject developing cancer (especially		
CC	breast cancer), or monitoring the effect of therapy on a subject with		
CC	cancer, comprising analysing a test sample using two-dimensional		
CC	electrophoresis and detecting Breast Cancer-Associated Protein		
CC	Isoforms (Bpis). The methods disclosed are used for the diagnosis and		
CC	prognosis of breast cancer, for determining the severity of breast		
CC	cancer, and for identifying the effect of therapy administered to a subject,		
CC	Antibodies raised against the binding domain of a BPI, the binding domain		
CC	of a BPI, a nucleic acid encoding a BPI, or a nucleic acid that inhibits		
CC	the function of a BPI can be incorporated into a pharmaceutical		
CC	composition for treating or preventing breast cancer. The methods use		
CC	sensitive and specific biomarkers provide early diagnosis of breast		
CC	cancer, and the compositions are more potent, specific, and has a more		
CC	rapid effect with fewer side effects than other prior art methods.		
CC	The present sequence is a tryptic digest peptide from a BPI of the		
XX	invention.		
SQ	Sequence 8 AA:		
QY	Query Match 100.0%; Score 20; DB 22; Length 8;		
Db	Best Local Similarity 100.0%; Pred. No. 6.4e+05;		
	Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0.		
	1 VNLD 4		
	3 VNLD 6		
RESULT 8			
AAEI0661	AAEI0661 standard; peptide; 8 AA.		
ID	AAEI0661		
AC	AAEI0661;		
XX			
DT	10-DEC-2001 (first entry)		
XX			
DE	Human aspartyl protease-1 beta-secretase Swedish mutant peptide.		
XX			
XX	Human; aspartyl protease 1; Asp1; amyloid precursor protein; APP;		
KW	Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;		
KW	amyloid plaque; neuronal loss; proteolytic; neuroprotective;		
KW	aspartyl protease-1 beta-secretase Swedish mutant peptide.		
XX			
OS	Homo sapiens.		
OS	Synthetic.		
XX			
FM	Key Location/Qualifiers		
FT	Cleavage-site 4..5		
XX			
PN	GB2357767-A.		
XX			
PD	04-JUL-2001.		
XX			
PF	22-SEP-2000; 2000GB-0023315.		
XX			
XX	23-SEP-1999; 99US-0155493.		
PR			

PR 23-SEP-1999; 99US-0404133.
 PR 23-SEP-1999; 99WO-US20881.
 PR 13-OCT-1999; 99US-0416901.
 PR 06-DEC-1999; 99US-0169232.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Blenkowski MJ, Gurney M;
 XX
 DR WPI: 2001-444208/48.
 XX
 PT Polypeptide comprising fragments of human aspartyl protease with
 PT amyloid precursor protein processing activity and alpha-secretase
 PT activity, for identifying modulators useful in treating Alzheimer's
 PT disease -
 XX
 PS Example 15; Page 92; 187pp; English.
 XX
 CC The patent discloses human aspartyl protease 1 (hu-Asp1) or modified
 CC Asp1 proteins which lack transmembrane domain or amino terminal
 CC domain or cytoplasmic domain and retains alpha-secretase activity
 CC and amyloid protein precursor (APP) processing activity. The proteins
 CC of the invention are useful for assaying hu-Asp1 alpha-secretase
 CC activity, which in turn is useful for identifying modulators of
 CC hu-Asp1 alpha-secretase activity, where modulators that increase
 CC hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's
 CC disease (AD) which causes progressive dementia with consequent
 CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
 CC neuronal loss. Hu-Asp1 protease substrate is useful for assaying
 CC hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with
 CC the substrate under acidic conditions and determining the level of
 CC hu-Asp1 proteolytic activity. The present sequence is human aspartyl
 CC protease-1 (hu-Asp-1) beta-secretase Swedish (Sw) mutant peptide
 CC which is used for determining the enzymatic activity of Asp-1 protein
 CC lacking a transmembrane (TM) domain and containing (His)6 tag.
 XX
 SQ Sequence 8 AA:
 Query Match 100.0%; Score 20; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VNLD 4
 ||||
 Db 2 VNLD 5
 RESULT 9
 AAE02613
 ID AAE02613 standard; peptide; 8 AA.
 XX
 AC AAE02613;
 XX
 DT 10-AUG-2001 (first entry)
 XX
 DE Human Aspartyl protease-1 beta-secretase Swedish mutant form peptide.
 XX
 KW Human; alpha-secretase; amyloid precursor protein; APP; therapy;
 KW Alzheimer's disease; antialzheimer's; aspartyl protease 1; Asp1;
 KW beta-secretase.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Cleavage-site 4..5
 XX
 PN WO200123533-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 22-SEP-2000; 2000WO-US26080.
 XX
 PR 23-SEP-1999; 99US-0155493.

PR 23-SEP-1999; 99WO-US20881.
 PR 13-OCT-1999; 99US-0416901.
 PR 06-DEC-1999; 99US-0169232.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Gurney M, Blenkowski MJ;
 XX
 DR WPI: 2001-290516/30.
 XX
 PT Enzymes that cleave the alpha-secretase site of the amyloid precursor
 PT protein, useful for the treatment of Alzheimer's disease -
 XX
 PS Example 15; Page 94; 189pp; English.
 XX
 CC The present invention relates to enzymes for cleaving the alpha-
 CC secretase site of the amyloid precursor protein (APP) and methods of
 CC identifying those enzymes. The methods may be used to identify enzymes
 CC that may be used to cleave the alpha-secretase cleavage site of the APP
 CC protein. The enzymes may be used to treat or modulate the progress of
 CC Alzheimer's disease. The present sequence is human aspartyl protease-1
 CC (hu-Asp-1) beta-secretase, Swedish (Sw) mutant form peptide which is used
 CC for determining the enzymatic activity of Asp-1 delatm (His)6 protein.
 XX
 SQ Sequence 8 AA:
 Query Match 100.0%; Score 20; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VNLD 4
 ||||
 Db 2 VNLD 5

RESULT 10
 AAW82085
 ID AAW82085 standard; peptide; 9 AA.
 XX
 AC AAW82085;
 XX
 DT 18-FEB-1999 (first entry)
 XX
 DE Fluorogenic protease indicator protease binding peptide #63.
 XX
 KW Protease activity; fluorophore; detection; fluorogenic; cellular uptake;
 KW conformation change.
 XX
 OS Synthetic.
 XX
 PN WO9837226-A1.
 XX
 PD 27-AUG-1998.
 XX
 PF 20-FEB-1998; 98WO-US03000.
 XX
 PR 20-FEB-1997; 97US-0802981.
 XX
 PA (ONCO-) ONCOIMMUNIN INC.
 XX
 PI Komoriya A, Packard BS;
 XX
 DR WPI: 1998-467579/40.
 XX
 PT New fluorogenic compositions - containing 2 fluorophores separated
 PT by a peptide comprising a protease binding site, used for detecting
 PT protease activity in samples.
 XX
 PS Claim 4; Page 77; 90pp; English.
 XX
 CC AAW82023-W82240 are peptides used in the construction of a fluorogenic
 CC composition which is used for the detection of protease activity in
 CC biological samples. The products can be used for the detection of

conformation changes in nucleic acids, oligosaccharides, polysaccharides, proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins, steroids or polymers. In addition, attachment of a hydrophobic group to a molecule can be used to enhance uptake by cells. The composition is composed of P = peptide comprising a protease binding site for the protease, F1, F2 peptides = fluorophores where F1 is attached to the amino terminal amino acid and F2 is attached to the carboxyl terminal amino acid and S1, S2 peptides = when present, are peptide spacers where S1, when present, is attached to the amino terminal amino acid, and S2, when present, is attached to the carboxyl terminal amino acid.

Sequence 9 AA;

Query Match 100.0%; Score 20; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
1111
DB 3 VNLD 6

RESULT 11

AAW82081
ID AAW82081 standard; peptide; 9 AA.

XX AAW82081;

XX 18-FEB-1999 (first entry)

XX Fluorogenic protease indicator protease binding peptide #59.

XX Protease activity; fluorophore; detection; fluorogenic; cellular uptake; conformation change.

XX Synthetic.

XX WO9837226-A1.

XX 27-AUG-1998.

XX 20-FEB-1998; 98WO-US03000.

XX 20-FEB-1997; 97US-0802981.

XX (ONCO-) ONCOIMMUNIN INC.

XX Komoriya A, Packard BS;

XX WPI: 1998-467579/40.

XX New fluorogenic compositions - containing 2 fluorophores separated by a peptide comprising a protease binding site, used for detecting protease activity in samples.

XX Claim 4; Page 77; 90pp; English.

XX AAW82023-W82240 are peptides used in the construction of a fluorogenic composition which is used for the detection of protease activity in biological samples. The products can be used for the detection of conformation changes in nucleic acids, oligosaccharides, polysaccharides, proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins, steroids or polymers. In addition, attachment of a hydrophobic group to a molecule can be used to enhance uptake by cells. The composition is composed of P = peptide comprising a protease binding site for the protease, F1, F2 peptides = fluorophores where F1 is attached to the amino terminal amino acid and F2 is attached to the carboxyl terminal amino acid and S1, S2 peptides = when present, are peptide spacers where S1, when present, is attached to the amino terminal amino acid, and S2, when present, is attached to the carboxyl terminal amino acid.

SQ Sequence 9 AA;

Query Match 100.0%; Score 20; DB 19; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
1111
DB 3 VNLD 6

RESULT 12

AAW07874
ID AAW07874 standard; peptide; 9 AA.

XX AAW07874;

XX 14-NOV-2000 (first entry)

XX A peptide fragment derived from beta-amyloid precursor protein.

XX Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide; amyloid plaque component; Alzheimer's disease; amyloidogenic disease;

XX inhibitor.

XX Homo sapiens.

XX WO200047618-A2.

XX 17-AUG-2000.

XX 10-FEB-2000; 2000WO-US03819.

XX 10-FEB-1999; 99US-0119571.

XX 15-JUN-1999; 99US-0139172.

XX (ELAN-) ELAN PHARM INC.

XX Anderson JP, Bast G, Doane MT, Frigon N, John V, Power M;

XX Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;

XX WPI: 2000-533011/48.

XX Purified beta-secretase protein used in assays to discover inhibitors which can be used for the treatment of amyloidogenic diseases e.g. Alzheimer's disease -

XX Disclosure; Page 12; 121pp; English.

XX The specification describes a beta-secretase enzyme. The enzyme cleaves beta-amyloid precursor protein to produce beta-amyloid peptide. This enzyme is therefore implicated in the production of amyloid plaque components which accumulate in the brains of individuals afflicted with CC Alzheimer's disease. Inhibitors of beta-secretase are administered to CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-like pathology to test if they maintain or improve cognitive CC ability or reduce the plaque burden. The compounds are used for the CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The CC present sequence represents a peptide derived from beta-amyloid CC precursor protein

XX Sequence 9 AA;

Query Match 100.0%; Score 20; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
1111
DB 3 VNLD 6

RESULT 13

AAB07894
ID AAB07894 standard; Peptide; 9 AA.
XX
AC AAB07894;
XX
DT 14-NOV-2000 (first entry)
XX
DE Substrate for beta-secretase enzyme.
XX
KM Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KM amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
XX inhibitor.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Cleavage-site 5..6
XX
PN WO20047618-A2.
XX
PD 17-AUG-2000.
XX
PF 10-FEB-2000; 2000WO-US03819.
XX
PR 10-FEB-1999; 99US-0119571.
XX
PR 15-JUN-1999; 99US-0139172.
XX
PA (ELAN) ELAN PHARM INC.
XX
PI Anderson JP, Basl G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX
DR WPI: 2000-533011/48.
XX
PT Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
PS Example 4: Page 71; 121pp; English.
XX
CC The specificity of the beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a peptide substrate used to test the
CC activity of beta-secretase enzyme.
XX
SQ Sequence 9 AA;
Query Match Best Local Similarity 100.0%; Score 20; DB 21; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VNLD 4
 ||||
DB 3 VNLD 6

RESULT 14
AAG73297
ID AAG73297 standard; Peptide; 9 AA.
XX
AC AAG73297;
XX
DT 14-AUG-2001 (first entry)
XX
DE Protease indicator compound peptide #26.
XX

KM Protease detection; peptide cleavage; enzyme activity; fluorogenic;
KM viral infection; cancer metastasis; emphysema; arthritis;
KM thrombosis; haemophilia.
XX
OS Synthetic.
XX
PN WO200118238-A1.
XX
PD 15-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-US24882.
XX
PR 10-SEP-1999; 99US-0394019.
XX
PA (ONCO-) ONCOMUNIN INC.
XX
PI Komoriya A, Packard BS;
XX
DR WPI: 2001-389573/41.
XX
PT New fluorogenic compositions whose fluorescence level increases in the
PT presence of active proteases, useful for detecting and localizing
PT protease activity in biological samples, particularly in frozen tissue
PT samples -
XX
PS Claim 1: Page 70; 86pp; English.
XX
CC The present invention describes fluorogenic compositions which can be
CC used for the detection of protease activity. This can be useful as an
CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,
CC thrombosis and arthritis. The fluorogenic compositions comprise a
CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The
CC peptide is cleaved by a protease and the fluorophores can then be
CC detected. The present sequence is one of the peptides described in the
CC exemplification of the invention.
XX
SQ Sequence 9 AA;
Query Match Best Local Similarity 100.0%; Score 20; DB 22; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VNLD 4
 ||||
DB 3 VNLD 6

RESULT 15
AAW08362
ID AAW08362 standard; peptide; 10 AA.
XX
AC AAW08362;
XX
DT 05-SEP-1997 (first entry)
XX
DE Beta-secretase substrate #3.
XX
KM Beta-cleavage site: beta amyloid precursor protein; APP; beta-secretase;
KM alpha-secretase; proteolytic cleavage; inhibitor; Alzheimer's disease.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note="acetylated"
XX
PN WO9640885-A2.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09985.
XX
PR 07-JUN-1995; 95US-0485152.

PR 07-JUN-1995; 9505-0480498.

XX (ATHE-) ATHENA NEUROSCIENCES INC.

XX Anderson JP, Chrysler SMS, Jacobson-croak KL, Keim PS;

PI Mcconlogue JC, Sinha S, Tan H;

XX WPI; 1997-052304/05.

DR Beta-secretase which specifically cleaves beta-amyloid precursor
XX protein - useful to screen for inhibitors useful in treatment of
PI Alzheimer's disease

PS Disclosure; Page 45; 92pp; English.

XX AAM08359-W08362 represent substrates for the enzyme of the invention.
CC The enzyme of the invention is beta-secretase, and specifically cleaves
CC beta-amyloid precursor protein (beta-APP). Normal processing of beta-APP
CC is thought to occur via cleavage between residues 16 and 17 of the
CC beta-amyloid peptide region by an alpha-secretase. Pathogenic processing
CC is thought to occur by beta-secretase cleavage of beta-APP.
CC Beta-secretase activity can be detected and measured using a method of
CC the invention, which detects at least one of the beta-secretase cleavage
CC products formed on cleavage. The method can be used to determine whether
CC a test substance inhibits proteolytic cleavage, by beta-secretase, of
CC beta-APP. Compounds effective to at least partially inhibit
CC beta-secretase activity can be used to inhibit cleavage of beta-APP in
CC cells or mammalian hosts. Isolation and purification of beta-secretase
CC will permit chemical modelling of a critical event in the pathology of
CC Alzheimer's disease.
XX

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 20; DB 18; Length 10;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4

Db 3 VNLD 6

Search completed: October 30, 2002, 12:27:11
Job time : 2.60197 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:24:20 ; Search time 0.570025 Seconds
(without alignments)
171.400 Million cell updates/sec

Title: US-09-724-571-104

Perfect score: 20
Sequence: 1 VNLD 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	US-09-054-334-6	Sequence 6, Appli
2	20	100.0	5	US-08-480-498-2	Sequence 2, Appli
3	20	100.0	5	US-08-659-984A-14	Sequence 14, Appli
4	20	100.0	5	US-08-660-531-14	Sequence 14, Appli
5	20	100.0	5	US-09-054-334-2	Sequence 2, Appli
6	20	100.0	9	US-08-802-981-219	Sequence 219, App
7	20	100.0	9	US-08-802-981-223	Sequence 223, App
8	20	100.0	10	US-08-659-984A-19	Sequence 19, Appli
9	20	100.0	10	US-08-660-531-19	Sequence 19, Appli
10	20	100.0	11	PCT-US94-07043A-3	Sequence 13, Appli
11	20	100.0	20	US-09-196-293-2	Sequence 2, Appli
12	20	100.0	20	US-08-209-603E-2	Sequence 2, Appli
13	20	100.0	21	US-08-659-984A-18	Sequence 18, Appli
14	20	100.0	21	US-08-596-257A-1	Sequence 1, Appli
15	20	100.0	21	US-08-802-981-112	Sequence 112, App
16	20	100.0	21	US-08-802-981-116	Sequence 116, App
17	20	100.0	21	US-08-860-339-1	Sequence 1, Appli
18	20	100.0	21	US-08-660-531-18	Sequence 18, Appli
19	20	100.0	30	US-08-659-984A-17	Sequence 17, Appli
20	20	100.0	30	US-08-660-531-17	Sequence 17, Appli
21	20	100.0	33	US-08-659-984A-16	Sequence 16, Appli
22	20	100.0	33	US-08-660-531-16	Sequence 16, Appli
23	20	100.0	42	US-08-659-984A-15	Sequence 15, Appli
24	20	100.0	42	US-08-660-531-15	Sequence 15, Appli
25	20	100.0	67	US-08-588-258B-6	Sequence 6, Appli
26	20	100.0	67	US-08-460-505-6	Sequence 6, Appli
27	20	100.0	67	PCT-US96-08295-6	Sequence 6, Appli

28	20	100.0	68	2	US-08-588-258B-10	Sequence 10, Appli
29	20	100.0	68	3	PCT-US96-08295-10	Sequence 10, Appli
30	20	100.0	68	5	PCT-US96-08295-10	Sequence 10, Appli
31	20	100.0	87	1	US-08-204-740-13	Sequence 13, Appli
32	20	100.0	87	3	US-09-081-167A-13	Sequence 13, Appli
33	20	100.0	87	3	US-09-081-395-13	Sequence 13, Appli
34	20	100.0	87	4	US-09-416-833-13	Sequence 13, Appli
35	20	100.0	87	5	PCT-US95-02521-13	Sequence 13, Appli
36	20	100.0	103	3	US-08-339-708A-12	Sequence 12, Appli
37	20	100.0	108	1	US-08-204-740-11	Sequence 11, Appli
38	20	100.0	108	3	US-09-081-167A-11	Sequence 11, Appli
39	20	100.0	108	3	US-09-081-395-11	Sequence 11, Appli
40	20	100.0	108	4	US-09-416-833-11	Sequence 11, Appli
41	20	100.0	108	5	PCT-US95-02521-11	Sequence 11, Appli
42	20	100.0	120	4	US-08-890-865A-13	Sequence 13, Appli
43	20	100.0	120	4	US-08-890-865A-14	Sequence 14, Appli
44	20	100.0	146	4	US-08-858-207A-494	Sequence 494, App
45	20	100.0	159	3	US-09-010-809-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-054-334-6
; Sequence 6, Application US/09054334
; Patent No. 6329163

GENERAL INFORMATION:

APPLICANT: Anderson, John P.
APPLICANT: Jacobson-Croak, Kirsten L.
APPLICANT: Sinha, Sukanto
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
TITLE OF INVENTION: Inhibition
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,334
FILING DATE: 02-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 015270-002820US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-054-334-6

Query Match 100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
1111
Db 1 VNLD 4

RESULT 2
US-08-480-498-2
Sequence 2, Application US/08480498
Patent No. 5744346
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Keim, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,498
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 015270-002200
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-480-498-2

Query Match 100.0%; Score 20; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
1111
Db 1 VNLD 4

RESULT 3
US-08-659-984A-14
Sequence 14, Application US/08659984A
Patent No. 5942400
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Sinha, Sukanto
APPLICANT: Jacobson-Croak, Kirsten L.
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
TITLE OF INVENTION: Inhibition
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California

COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,984A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-659-984A-14

Query Match 100.0%; Score 20; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
1111
Db 1 VNLD 4

RESULT 4
US-08-660-531-14
Sequence 14, Application US/08660531
Patent No. 6221645
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Keim, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,531
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,498
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-660-531-14

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 5;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 1 VNLD 4

RESULT 5
US-09-054-334-2
Sequence 2, Application US/09054334
Patent No. 6329163
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Jacobson-Croak, Kirsten L.
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
TITLE OF INVENTION: Inhibition
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,334
FILING DATE: 02-APR-1998
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 015270-002820US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-054-334-2

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 5;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VNLD 4
||||

DB 1 VNLD 4

RESULT 6
US-08-802-981-219
Sequence 219, Application US/08802981
Patent No. 6037137
GENERAL INFORMATION:
APPLICANT: Komoriya, Akira
APPLICANT: Packard, Beverly S.
TITLE OF INVENTION: Compositions for the Detection of Enzyme
TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/802,981
FILING DATE: 20-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 016865-000300US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-802-981-219

Query Match
Best Local Similarity 100.0%; Score 20; DB 3; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 3 VNLD 6

RESULT 7
US-08-802-981-223
Sequence 223, Application US/08802981
Patent No. 6037137
GENERAL INFORMATION:
APPLICANT: Komoriya, Akira
APPLICANT: Packard, Beverly S.
TITLE OF INVENTION: Compositions for the Detection of Enzyme
TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/802,981
FILING DATE: 20-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 016865-000300US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 223:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-802-981-223

Query Match
Best Local Similarity 100.0%; Score 20; DB 3; Length 9;
Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
1111
DB 3 VNLD 6

RESULT 8
US-08-659-984A-19
Sequence 19, Application US/08659984A
Patent No. 5942400
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Sinha, Sukanto
APPLICANT: Jacobson-Croak, Kirsten L.
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
TITLE OF INVENTION: Inhibition
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,984A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2422
TELEFAX: 415-326-2400
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: one-of(1)
OTHER INFORMATION: /note= "N-terminal Ser is acetylated."
US-08-659-984A-19

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 10;
Pred. No. 23;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
1111
DB 3 VNLD 6

RESULT 9
US-08-660-531-19
Sequence 19, Application US/08660531
Patent No. 6221645
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Keim, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,531
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,498
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: one-of(1)
OTHER INFORMATION: /note= "N-terminal Ser is acetylated."
US-08-660-531-19

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 10;
Pred. No. 23;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4

||||
DB 3 VNLD 6

RESULT 10

PCT-US94-07043A-3
Sequence 3, Application PC/TUS9407043A

GENERAL INFORMATION:

APPLICANT: Tamburini, Paul P.; Benz, G nter; H bich,
APPLICANT: Dieter; Dreyer, Robert N.; Koenig, Gerhard
TITLE OF INVENTION: CATHEPSIN D IS AN AMYLOIDGENIC
TITLE OF INVENTION: PROTEASE IN ALZHEIMER S DISEASE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Miles Inc.
STREET: 400 Morgan Lane
CITY: West Haven

STATE: Connecticut
COUNTRY: USA
ZIP: 06516

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage
COMPUTER: Sharp PC 4600
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/07043A
FILING DATE: June 21, 1994
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/10889
FILING DATE: November 12, 1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/995,660
FILING DATE: December 16, 1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/880,914
FILING DATE: May 11, 1992

ATTORNEY/AGENT INFORMATION:

NAME: Pamela A. Simonson
REGISTRATION NUMBER: 31,060

REFERENCE/DOCKET NUMBER: MTI 224.3
TELECOMMUNICATION INFORMATION:

TELEPHONE: (203) 937-2340
TELEFAX: (203) 937-2795

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: linear

PCT-US94-07043A-3

Query Match 100.0%; Score 20; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 26;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 4 VNLD 7

RESULT 11

US-09-196-293-2
Sequence 2, Application US/09196293

GENERAL INFORMATION:

APPLICANT: Fuchs, Renate
APPLICANT: Motz, Manfred
APPLICANT: Sautscheck, Erwin

APPLICANT: Wilske, Bettina
APPLICANT: Preac-Mursic, Vera

TITLE OF INVENTION: Active proteins from Borrelia
TITLE OF INVENTION: burgdorferi

FILE REFERENCE: 738 001US2
CURRENT APPLICATION NUMBER: US/09/196,293
CURRENT FILING DATE: 1998-11-19
EARLIER APPLICATION NUMBER: US 08/209,603
EARLIER FILING DATE: 1994-03-10
EARLIER APPLICATION NUMBER: US 07/862,535
EARLIER FILING DATE: 1992-06-19
EARLIER APPLICATION NUMBER: WO PCT/EP90/02282
EARLIER FILING DATE: 1990-12-21
EARLIER APPLICATION NUMBER: DE P39 42 728.5
EARLIER FILING DATE: 1989-12-22
EARLIER APPLICATION NUMBER: DE P40 18 988.0
EARLIER FILING DATE: 1990-06-13
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 20
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-196-293-2

Query Match 100.0%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 11 VNLD 14

RESULT 12

US-08-209-603E-2
Sequence 2, Application US/08209603E

GENERAL INFORMATION:

APPLICANT: FUCHS, RENATE
APPLICANT: WILSKE, BETTINA
APPLICANT: PREAC-MURISIC, VERA

APPLICANT: MOTZ, MANFRED
APPLICANT: SAUTSCHECK, ERWIN

TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PROTEINS
TITLE OF INVENTION: FROM BORRELLIA BURGDORFERI
NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK

STATE: NY
COUNTRY: USA
ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2

SOFTWARE: ASCII
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/209,603E
FILING DATE: 10-MAR-1994
CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP90/02282
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/862,535

FILING DATE: 19-JUN-1992
ATTORNEY/AGENT INFORMATION:

NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224

REFERENCE/DOCKET NUMBER: LKR-9217-A
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 20
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: N/A
ORIGINAL SOURCE: B. BURGDOFFER
IMMEDIATE SOURCE: DSM 5662
LIBRARY: N/A
POSITION IN GENOME: N/A
FEATURE:
IDENTIFICATION METHOD: amino acid analysis
PUBLICATION INFORMATION: N/A
US-08-209-603E-2

Query Match 100.0%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
DB 11 VNLD 14

RESULT 13
US-08-659-984A-18
; Sequence 18, Application US/08659984A
; Patent No. 5942400
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Sinha, Sukanto
; APPLICANT: Jacobson-Croak, Kirsten L.
; TITLE OF INVENTION: Assays for Detecting Beta-Secretase
; TITLE OF INVENTION: Inhibition
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Ctr., 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,984A
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,152
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-002810US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-659-984A-18

Query Match 100.0%; Score 20; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
DB 3 VNLD 6

RESULT 14
US-08-596-257A-1
; Sequence 1, Application US/08596257A
; Patent No. 6001628
; GENERAL INFORMATION:
; APPLICANT: KOSSMAN, Jens
; APPLICANT: EMMERMAN, Michael
; APPLICANT: VIRGIN, Ivair
; TITLE OF INVENTION: BRANCHING ENZYMES AND DNA SEQUENCES
; TITLE OF INVENTION: CODING THEM, SUITABLE FOR CHANGING THE DEGREE OF BRANCHING
; TITLE OF INVENTION: OF AMYLOPECTIN STARCH IN PLANTS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/596,257A
; FILING DATE: 08-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4327165.0
; FILING DATE: 09-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meilman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/951-118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-596-257A-1

Query Match 100.0%; Score 20; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
DB 5 VNLD 8

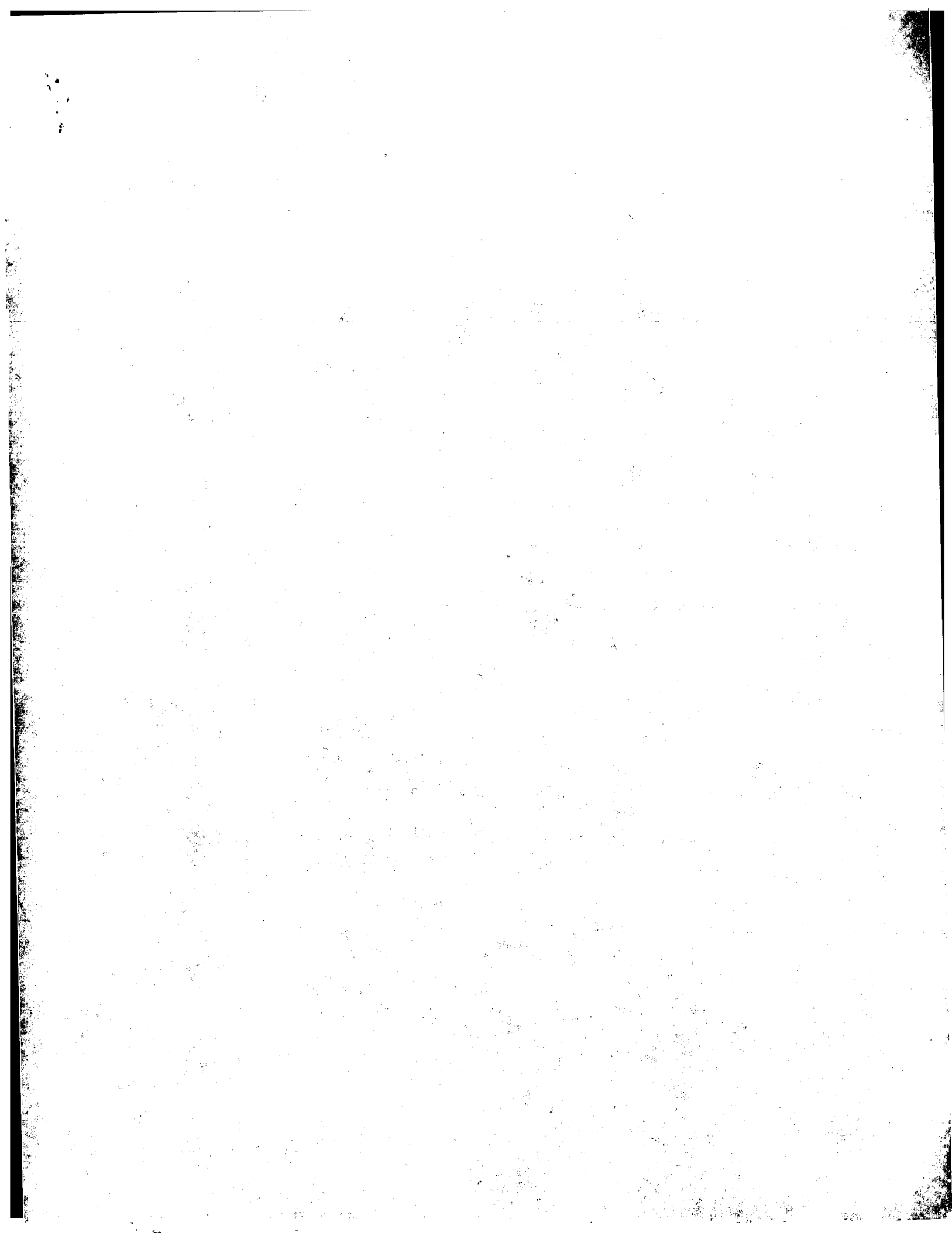
RESULT 15
US-08-802-981-112
; Sequence 112, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.

TITLE OF INVENTION: Compositions for the Detection of Enzyme
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/802,981
FILING DATE: 20-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 016865-000300US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /product= "Alb"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /product= "Acp"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 16
OTHER INFORMATION: /product= "Acp"
US-08-802-981-112

Query Match 100.0%; Score 20; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
Db 8 VNLD 11

Search completed: October 30, 2002, 12:32:37
Job time : 1.57002 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:23:30 ; Search time 0.766585 Seconds

(without alignments)
501.389 Million cell updates/sec

Title: US-09-724-571-104

Perfect score: 20
Sequence: 1 VNLD 4

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	20	100.0	15	2 P00017	terminal protein -
2	20	100.0	30	2 S27306	surface layer prot
3	20	100.0	33	2 H95094	hypothetical prote
4	20	100.0	33	2 E82852	hypothetical prote
5	20	100.0	49	2 T07309	hypothetical prote
6	20	100.0	59	2 T38172	hypothetical prote
7	20	100.0	62	2 S45295	hypothetical prote
8	20	100.0	62	2 S27308	cyclin Cyc3 - Arab
9	20	100.0	63	2 S04973	surface layer prot
10	20	100.0	72	2 S34404	antifreeze protein
11	20	100.0	79	2 T17694	ribosomal protein
12	20	100.0	80	2 H89990	hypothetical prote
13	20	100.0	83	2 T36483	hypothetical prote
14	20	100.0	84	2 E70841	probable membrane
15	20	100.0	88	2 S70644	eclosion hormone p
16	20	100.0	89	2 S54814	probable acyl car
17	20	100.0	89	2 H87367	hypothetical prote
18	20	100.0	90	2 T46007	hypothetical prote
19	20	100.0	91	2 S57268	translation elonga
20	20	100.0	91	2 T29686	hypothetical prote
21	20	100.0	91	2 A75188	translation elonga
22	20	100.0	91	2 A11583	hypothetical prote
23	20	100.0	93	2 G81410	hydrogenase isoenz
24	20	100.0	96	2 PC4086	ribosomal protein
25	20	100.0	97	2 S2857	hypothetical prote
26	20	100.0	99	2 T06983	root abundant prot
27	20	100.0	104	2 A69407	signal recognition
28	20	100.0	104	2 E86263	flk23.22 protein
29	20	100.0	105	2 T42021	Ser/Thr protein ph

30	20	100.0	106	2 A10113	probable Rieske pr
31	20	100.0	106	2 JG0021	flagellar hook-bas
32	20	100.0	109	2 B70637	hypothetical prote
33	20	100.0	110	2 S65003	hypothetical prote
34	20	100.0	112	2 A69451	conserved hypotnet
35	20	100.0	112	2 B95061	BLPS protein limpo
36	20	100.0	112	2 P97929	regulatory protein
37	20	100.0	115	2 T48186	light-inducible pr
38	20	100.0	116	1 NGN1X1	nerve growth facto
39	20	100.0	116	2 A58566	N-acetylmuramoyl-L
40	20	100.0	116	2 G64126	conserved hypotnet
41	20	100.0	117	2 F64708	holo-(acyl carrier
42	20	100.0	119	2 P83714	probable sensory t
43	20	100.0	120	2 B95269	ribosomal protein
44	20	100.0	121	2 S78265	hypothetical prote
45	20	100.0	123	2 C81004	

ALIGNMENTS

RESULT 1
P00017
terminal protein - phage M2 (fragment)
C:Species: phage M2
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C:Accession: P00017
R:Matsumoto, K.; Takano, H.; Kim, C.I.; Hirokawa, H.
Gene 84, 247-255, 1989
A>Title: Primary structure of bacteriophage M2 DNA polymerase: conserved segments wit
A:Reference number: J00161; MUID:90128268
A:Accession: P00017
A:Molecule type: DNA
A:Residues: 1-15 <MAT>
A:Cross-references: GB:M33144; NID:g215507; PIDN:AAA32367.1; PID:g215508
C:Genetics:
A:Gene: E
C:Superfamily: phage PZA terminal protein

Query Match 100.0%; Score 20; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
|||
Db 8 VNLD 11

RESULT 2

S27306
surface layer protein - Aeromonas hydrophila
N:Alternate names: S-layer protein
C:Species: Aeromonas hydrophila
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
R:Dooley, J.S.G.; McCubbin, W.D.; Kay, C.M.; Trust, T.J.
J. Bacteriol. 170, 2631-2638, 1988
A>Title: Isolation and biochemical characterization of the S-layer protein from a pat
A:Reference number: S27306; MUID:88227842
A:Accession: S27306
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-30 <DOO>

Query Match 100.0%; Score 20; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
|||
Db 1 VNLD 4

RESULT 3
H95094
hypothetical protein SP0821 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: H95094
R:Retelien, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
non, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95094
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-33 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK74953.1; PID:914972294; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0821

Query Match 100.0%; Score 20; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
|||
Db 5 VNLD 8

RESULT 4
E82852
hypothetical protein XF0065 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: E82852
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-33 <STM>
A:Cross-references: GB:AE003860; GB:AE003849; NID:9104830; PIDN:AAF82878.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Brienes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
Rodrigues, V.; Rosa, A.O. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0065

Query Match 100.0%; Score 20; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
|||
Db 15 VNLD 18

RESULT 5
T07309
hypothetical protein 49e - Chlorella vulgaris chloroplast
C:Species: Chlorella vulgaris
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C:Accession: T07309
R:Wakagagi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga C
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A:Reference number: Z15985; MUID:97303241
A:Accession: T07309
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-49 <NAK>
A:Cross-references: EMBL:AB001684; NID:92224352; PIDN:BA57957.1; PID:92224473
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 100.0%; Score 20; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 1,3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
|||
Db 3 VNLD 6

RESULT 6
T38172
hypothetical protein SPAC22E12.15 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38172
R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z21775
A:Accession: T38172
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-59 <DEV>
A:Cross-references: EMBL:Z70043; NID:91220275; PIDN:CAA93902.1; GSPDB:GN00066; SPDB:S
A:Experimental source: strain 972h; cosmid c22E12
C:Genetics:
A:Gene: SPDB:SPAC22E12.15
A:Map position: 1
A:introns: 19/1; 23/1

Query Match 100.0%; Score 20; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
|||
Db 19 VNLD 22

RESULT 7
S45295
cyclin Cyc3 - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 13-Jan-1995 #sequence_revision 19-Apr-1996 #text_change 16-Jul-1999
C:Accession: S45295
R:Day, I.S.; Reddy, A.S.N.
Biochim. Biophys. Acta 1218, 115-118, 1994
A:Title: Cloning of a family of cyclins from Arabidopsis thaliana.
A:Reference number: S45293; MUID:94250688
A:Accession: S45295
A:Molecule type: mRNA
A:Residues: 1-62 <DAY>
A:Cross-references: EMBL:L27225; NID:9456021; PIDN:AAA19880.1; PID:9456022
C:Superfamily: cyclin

C:Keywords: cell cycle control

Query Match 100.0%; Score 20; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 27 VNLD 30

RESULT 8
S27308

Surface layer protein - Aeromonas hydrophila
N:Alternate names: S-layer protein
C:Species: Aeromonas hydrophila

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-1999
C:Accession: S27308
R:Kosarzynska, M.; Dooley, J.S.G.; Shimojo, T.; Sakata, T.; Trust, T.J.
J. Bacteriol. 174, 40-47, 1992

A:Title: Antigenic diversity of the S-layer proteins from pathogenic strains of Aeromonas
A:Reference number: S27308; MUID:92104988
A:Accession: S27308

A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-62 <MOS>

Query Match 100.0%; Score 20; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 1 VNLD 4

RESULT 9
S04973

Antifreeze protein ABI - Antarctic eelpout
C:Species: Austrolyctichthys brachycephalus (Antarctic eelpout)

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 24-Oct-2000
C:Accession: S04973; B30095
R:Cheng, C.H.C.; Devries, A.L.
Biochim. Biophys. Acta 997, 55-64, 1989

A:Title: Structures of antifreeze peptides from the antarctic eel pout, Austrolyctichthys
A:Reference number: S04973; MUID:89323219
A:Accession: S04973
A:Molecule type: protein

A:Residues: 1-63 <CHE>
A>Note: the accession number A30095 cited as assigned by EMBL is not in GenBank release

C:Superfamily: antifreeze protein SPI
C:Keywords: antifreeze

Query Match 100.0%; Score 20; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 48 VNLD 51

RESULT 10
S34404

ribosomal protein L15 - Bacillus licheniformis (fragment)
C:Species: Bacillus licheniformis

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-1999
C:Accession: S34404
R:Tschauder, S.; Driessen, A.J.M.; Freudl, R.
Mol. Gen. Genet. 235, 147-152, 1992

A:Title: Cloning and molecular characterization of the secY genes from Bacillus licheniformis
A:Reference number: S30115; MUID:93062802
A:Accession: S34404

A:Molecule type: DNA
A:Residues: 1-72 <TSC>
A:Cross-references: EMBL:X70087
C:Genetics:
A:Gene: rpl15
C:Superfamily: Escherichia coli ribosomal protein L15
C:Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 20; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 3 VNLD 6

RESULT 11
T17694

hypothetical protein a204L - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17694
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999

A:Reference number: T17694
A:Accession: T17694

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-79 <GRA>

A:Cross-references: EMBL:U42580; NID:94028896; PIDN:AMC96572.1
A:Experimental source: specific host Chlorella strain NC6A

C:Genetics:
A>Note: a204L

Query Match 100.0%; Score 20; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 48 VNLD 51

RESULT 12
H89990

hypothetical protein SA1809 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: H89990
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O

ma, A.; Mizutani-Tl, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C. Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H89990

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-80 <KUR>
A:Cross-references: GB:BA000018; PID:G13701796; PIDN:BA843089.1; GSPDB:GN00149

A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1809

Query Match 100.0%; Score 20; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
||||
DB 67 VNLD 70

RESULT 13

T36483

probable membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: J36483

R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1999

A:Reference number: 221608

A:Accession: J36483

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-83 <SAD>

A:Cross-references: EMBL:AL096822; PIDN:CAB46925.1; GSPDB:GN00070; SCOEDB:SCGD3.03

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SCGD3.03

Query Match

Best Local Similarity 100.0%; Score 20; DB 2; Length 83;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4

DB 27 VNLD 30

RESULT 14

E70841

hypothetical protein RV2104c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: E70841

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987

A:Accession: E70841

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-84 <COL>

A:Cross-references: GB:AL021924; GB:AL123456; NID:g3261519; PIDN:CA17287.1; PID:el25238

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV2104c

Query Match

Best Local Similarity 100.0%; Score 20; DB 2; Length 84;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4

DB 59 VNLD 62

RESULT 15

JS0644

eclosion hormone precursor - silkworm

C:Species: Bombyx mori (silkworm)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Jul-2000

C:Accession: JS0644

R:Kimoto, T.; Tanaka, H.; Sato, B.; Nagasawa, H.; Suzuki, A.

Biochem. Biophys. Res. Commun. 182, 514-519, 1992

A:Title: Nucleotide sequence of cDNA for the eclosion hormone of the silkworm, Bombyx mori

A:Reference number: JS0644; MUID:92134263

A:Accession: JS0644

A:Molecule type: mRNA

A:Residues: 1-88 <KAM>

A:Cross-references: GB:D10135; NID:g217271; PIDN:BAA01012.1; PID:g217272

C:Comment: This neuropeptide controls ecdysis.

C:Superfamily: eclosion hormone

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-88/Product: eclosion hormone #status predicted <ECL>

Query Match

Best Local Similarity 100.0%; Score 20; DB 2; Length 88;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4

DB 19 VNLD 22

Search completed: October 30, 2002, 12:31:41
Job time : 3.7658 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:10 ; Search time 0.34398 Seconds

(Without alignments)
450.253 Million cell updates/sec

Title: US-09-724-571-104

Perfect score: 20

Sequence: 1 VNLD 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	15	1	TERM_BPM2
2	20	100.0	40	1	CSP4_STRMT
3	20	100.0	59	1	YDBF-SCHRO
4	20	100.0	63	1	ANP1-AUSBR
5	20	100.0	67	1	RGS3_RAT
6	20	100.0	72	1	RL15_BACLI
7	20	100.0	88	1	ECLH_BOMMO
8	20	100.0	90	1	EP1B-SURSO
9	20	100.0	91	1	EP1B-PYRAB
10	20	100.0	96	1	RL15-STRSC
11	20	100.0	104	1	SR19-ARCFU
12	20	100.0	105	1	HAL2-CANAL
13	20	100.0	106	1	FLIE_BACSU
14	20	100.0	112	1	SPEH-ARCFU
15	20	100.0	116	1	NGF_NAJAR
16	20	100.0	116	1	NGF_NAJAR
17	20	100.0	116	1	YEP4_HAETN
18	20	100.0	119	1	ACPS_BACHD
19	20	100.0	121	1	RK14-ODOSI
20	20	100.0	125	1	DHSC-COXHU
21	20	100.0	129	1	DHSC-ECOLI
22	20	100.0	131	1	FABA-BOVIN
23	20	100.0	145	1	YPH2_MYCCA
24	20	100.0	146	1	R27B-ARATH
25	20	100.0	146	1	R27C-ARATH
26	20	100.0	146	1	RL15_BACSU
27	20	100.0	146	1	RL15-STAM
28	20	100.0	147	1	RL2A-HUMAN
29	20	100.0	147	1	RL2A-MOUSE
30	20	100.0	147	1	RL2A-RAT
31	20	100.0	148	1	R28A-SCHPO
32	20	100.0	148	1	R28B-SCHPO
33	20	100.0	150	1	RL15-MICLU

34	20	100.0	151	1	RL15-STRCO	P46787 streptomyc
35	20	100.0	151	1	TCPR_VIBCH	P29483 vibrio chol
36	20	100.0	151	1	YABR_HAEIN	P45056 haemophilus
37	20	100.0	152	1	TR13_HUMAN	Q15649 homo sapien
38	20	100.0	152	1	YABR-ECOLI	P22186 escherichia
39	20	100.0	153	1	NRDI_MYCPN	P43961 mycoplasma
40	20	100.0	160	1	Y178_HAEIN	P43961 haemophilus
41	20	100.0	166	1	IL3_RAT	P04823 rattus norv
42	20	100.0	168	1	YWV1_CAEEL	Q11088 caenorhabdi
43	20	100.0	170	1	YFIR_BACHD	O94867 bacillus ha
44	20	100.0	172	1	YFIR-ECOLI	P76597 escherichia
45	20	100.0	173	1	DEF_BUCAT	P57563 buchnera ap

ALIGNMENTS

RESULT 1
ID TERM_BPM2 STANDARD: PRT: 15 AA.
AC P19897;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA terminal protein (Protein GP3) (Fragment).
GN 3 OR F.
OS Bacteriophage M2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.
OX NCBI_TaxID=10751;
RN [1]
RP MEDLINE=90128268; PubMed=2515115;
RA Matsumoto K., Takano H., Kim C.I., Hirokawa H.;

"Primary structure of bacteriophage M2 DNA polymerase: conserved segments within protein-priming DNA polymerases and DNA polymerase I or Escherichia coli.";
RL Gene 84:247-255(1989).
CC -!- FUNCTION: DNA TERMINAL PROTEIN IS LINKED TO THE 5' ENDS OF BOTH STRANDS OF THE GENOME THROUGH A PHOSPHODIESTER BOND BETWEEN THE BETA-HYDROXYL GROUP OF A SERINE RESIDUE AND THE 5'-PHOSPHATE OF THE TERMINAL DEOXYADENYLATE. THIS PROTEIN IS ESSENTIAL FOR DNA REPLICATION AND IS INVOLVED IN THE PRIMING OF DNA ELONGATION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: M33144; AAA32367.1; -.
DR PIR: PQ0017; PQ0017.
KW Early protein; DNA replication; DNA priming;
KW Covalent protein-DNA linkage.
FT NON-TER 1
FT SITE 5 7
FT SEQUENCE 15 AA: 1797 MW: D3CBAF8759DEA06 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VNLD 4
|||
Db 8 VNLD 11

RESULT 2
CSP4_STRMT STANDARD: PRT: 40 AA.
ID CSP4_STRMT
AC O33675;
DT 30-MAY-2000 (Rel. 39, Created)

```

DT 30-MAY-2000 (Rel. 39, last sequence update)
DE 30-MAY-2000 (Rel. 39, last annotation update)
DE Competence stimulating peptide precursor (CSP).
GN COMC.
OS Streptococcus mlts.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=28037;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NTC 12261;
RX MEDLINE=98012953; PubMed=9352904;
RX Haeverstein L.S., Hakenbeck R., Gaustad P.;
RT "Natural competence in the genus Streptococcus: evidence that
RT streptococci can change phenotype by interspecies recombinational
RT exchanges."
RL J. Bacteriol. 179:6589-6594(1997).
CC -i- FUNCTION: ACTS AS A PHEROMONE, INDUCES CELLS TO DEVELOP COMPETENCE
CC FOR GENETIC TRANSFORMATION.
CC -i- SUBCELLULAR LOCATION: Secreted.
CC -i- SIMILARITY: BELONGS TO THE COMC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ000875; CA04365.1; -
DR InterPro: IPR004288; COMC.
DR Pfam: PF03047; COMC. 1.
KM Pheromone; Competence.
FT PROPEP 1 24
FT CHAIN 25 40
FT SEQUENCE 40 AA; 4898 MW; 599DC7A3BADA1822 CRC64;
SQ
Query Match 100.0%; Score 20; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
DB 5 VNLD 8

RESULT 3
YDBF_SCHPO
ID YDBF_SCHPO STANDARD: PRT; 59 AA.
AC 010365;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, last sequence update)
DE 01-OCT-1996 (Rel. 34, last annotation update)
DE Hypothetical 6.6 kDa protein C22E12.15 in chromosome I.
GN SPAC22E12.15.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RT Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z70043; CA93902.1; -
DR Hypothetical protein.
KM SEQUENCE 59 AA; 6594 MW; 39D9EAO04168549 CRC64;
SQ
Query Match 100.0%; Score 20; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
DB 19 VNLD 22

RESULT 4
ANPL_AUSBR
ID ANPL_AUSBR STANDARD: PRT; 63 AA.
AC P12100;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, last sequence update)
DT 15-DEC-1998 (Rel. 37, last annotation update)
DE Antifreeze peptide ABL.
OS Austrololichthys brachycephalus (Antarctic eel pout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Zoarcoidei;
OC Zoarcidae; Austrololichthys.
OX NCBI_TaxID=8195;
RN [1]
RP SEQUENCE.
RX MEDLINE=89323219; PubMed=2752054;
RX Cheng C.-H.C., Devries A.L.;
RT "Structures of antifreeze peptides from the antarctic eel pout,
RT Austrololichthys brachycephalus."
RT Biochim. Biophys. Acta 937:55-64(1989).
CC -i- FUNCTION: ANTIFREEZE PROTEINS LOWER THE BLOOD FREEZING POINT BY
CC ABSORBING ICE AND INHIBITING ITS GROWTH.
CC -i- SIMILARITY: BELONGS TO THE TYPE-III AFP FAMILY.
CC -----
DR PIR: S04973; S04973.
DR HSP: P35753; 3RDN.
DR InterPro: IPR000029; Antifreeze_type_III.
DR Pfam: PF01354; Antifreeze_1.
DR PRINTS: PR00357; ANTIFREEZE211.
DR ProDom: PD003258; Antifreeze_type_III; 1.
KM Antifreeze protein.
FT SITE 8 8 IMPORTANT FOR ICE-BINDING
FT SITE 13 13 (BY SIMILARITY).
FT SITE 17 17 IMPORTANT FOR ICE-BINDING
FT SITE 43 43 (BY SIMILARITY).
FT SITE 43 43 IMPORTANT FOR ICE-BINDING
FT SITE 43 43 (BY SIMILARITY).
SQ SEQUENCE 63 AA; 6846 MW; AC84FD1424719384 CRC64;
Query Match 100.0%; Score 20; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
DB 48 VNLD 51

RESULT 5
RGS3_RAT
ID RGS3_RAT STANDARD: PRT; 67 AA.
AC P49797;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, last sequence update)
DT 01-MAR-2002 (Rel. 41, last annotation update)
DE Regulator of G-protein signaling 3 (RGS3) (Fragment).

```

```

GN RGS3.
ON Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=96140645; PubMed=8548815;
RA Koelle M.R., Horvitz H.R.;
RT "Egr-10 regulates G protein signaling in the C. elegans nervous
RL system and shares a conserved domain with many mammalian proteins."
RL Cell 84:115-125(1996).
RN [2]
RP PHOSPHORYLATION.
RX MEDLINE=20167219; PubMed=10702309;
RA Pedram A., Razandi M., Kehrl J., Levin E.R.;
RT "Natriuretic peptides inhibit G protein activation. Mediation through
RT cross-talk between cyclic GMP-dependent protein kinase and regulators
RT of G protein-signaling proteins."
RL J. Biol. Chem. 275:7365-7372(2000).
CC -1- FUNCTION: INHIBITS SIGNAL TRANSDUCTION BY INCREASING THE GTPASE
CC ACTIVITY OF G PROTEIN ALPHA SUBUNITS THEREBY DRIVING THEM INTO
CC THEIR INACTIVE GDP-BOUND FORM.
CC -1- PHOSPHORYLATED BY CYCLIC GMP-DEPENDENT PROTEIN KINASE.
CC -1- SIMILARITY: CONTAINS 1 RGS DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U32434; AAC52371.1; -.
DR HSRP: P49799; IAGR.
DR InterPro: IPR000342; RGS.
DR Pfam: PF00615; RGS; 1.
DR ProDom: PD001580; RGS; 1.
DR PROSITE: PS00132; RGS; 1.
KW Signal transduction inhibitor; Phosphorylation.
FT NON_TER 1
FT DOMAIN 1
FT DON_TER <1 >67 RGS.
FT NON_TER 67
FT SEQUENCE 67 AA; 7641 MW; 06FE7630E1CE3AF5 CRC64;
SQ
Query Match 100.0%; Score 20; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VMLD 4
Db 35 VMLD 38

```

```

RT RT "Cloning and molecular characterization of the secy genes from
RT Bacillus licheniformis and Staphylococcus carnosus: comparative
RT analysis of nine members of the Secy family.";
RL Mol. Gen. Genet. 235:147-152(1992).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -1- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X70087; -; NOT_ANNOTATED_CDS.
DR PIR: S34404; S34404.
DR InterPro: IPR001196; Ribosomal_L15.
DR Pfam: PF00256; L15; 1.
DR PROSITE: PS00475; RIBOSOMAL_L15; 1.
KW Ribosomal protein; rRNA-binding.
FT NON_TER 1
FT SEQUENCE 72 AA; 7441 MW; 6DC3BA681492E6C9 CRC64;
SQ
Query Match 100.0%; Score 20; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VMLD 4
Db 3 VMLD 6

```

RESULT 7

```

ECLH_BOMMO
ID ECLH_BOMMO STANDARD; PRT; 88 AA.
AC P25331;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Eclosion hormone precursor (Ecdisis activator) (EH).
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92134263; PubMed=1370883;
RA Kamito T., Tanaka H., Sato B., Nagasawa H., Suzuki A.;
RT "Nucleotide sequence of cDNA for the eclosion hormone of the
RL silkworm, Bombyx mori, and the expression in a brain.";
RL Biochem. Biophys. Res. Commun. 182:514-519(1992).
RN [2]
RP SEQUENCE OF 27-87.
RA Kono T., Nagasawa H., Isogai A., Fugo H., Suzuki A.;
RT "Amino acid sequence of eclosion hormone of the silkworm, Bombyx
RL mori.";
RL Agric. Biol. Chem. 51:2307-2308(1987).
CC -1- FUNCTION: NEUROPEPTIDE THAT TRIGGERS THE PERFORMANCE OF ECDYSIS
CC BEHAVIORS AT THE END OF A MOLT. IT TRIGGERS ADULT BEHAVIOR
CC PATTERNS: LARVAL, PUPAL AND ADULT ECDYSIS, AND PLASTICIZATION
CC DURING THE MOLT.
CC -1- SIMILARITY: HIGH, TO OTHER INSECTS ECLOSION HORMONE.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----

```

```

DR EMBL: D10135; BAA01012.1; -.
DR PIR: JS0644; JS0644.
KW Hormone; Neuropeptide; Signal.
FT SIGNAL 1 26
FT CHAIN 27 88 ECLOSION HORMONE.
FT DISULFID 40 64 BY SIMILARITY.
FT DISULFID 44 60 BY SIMILARITY.
FT DISULFID 47 75 BY SIMILARITY.
FT DISULFID 31 31 A -> S (IN REF. 2).
FT CONFLICT 33 33 S -> G (IN REF. 2).
FT CONFLICT 46 46 Q -> E (IN REF. 2).
FT CONFLICT 67 68 AR -> FK (IN REF. 2).
FT CONFLICT 70 71 KD -> DL (IN REF. 2).
FT CONFLICT 71 71
SQ SEQUENCE 88 AA; 9505 MW; 429DDCC0ADDAAD52 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 UNLD 4
DB 19 UNLD 22

RESULT 8
EF1B_SULSO STANDARD; PRT; 90 AA.
ID EF1B_SULSO
AC Q64214;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor 1-beta (EF-1-beta) (aEF-1beta).
GN EF1B OR S505345.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-11 AND 82-90.
RC STRAIN=DSM 5833 / Mt-4;
RX MEDLINE=95359209; PubMed=7632739;
RA Arcari P., Kaimo G., Iannicello G., Gallo M., Bocchini V.;
RT "The first nucleotide sequence of an archaeal elongation factor 1
beta gene.";
RL Biochim. Biophys. Acta 1263:86-88(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyer M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtlis B.A.,
RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
RN [3]
RP SEQUENCE, FUNCTION, SUBUNITS, AND MASS SPECTROMETRY.
RC STRAIN=DSM 5833 / Mt-4;
RX MEDLINE=96186282; PubMed=8652615;
RA Raimo G., Masullo M., Savino G., Searano G., Iannicello G.,
RA Parente A., Bocchini V.;
RT "Archaeal elongation factor 1 beta is a dimer. Primary structure,
molecular and biochemical properties.";
RL Biochim. Biophys. Acta 1293:106-112(1996).
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF GDP FOR GTP IN EF-1-ALPHA/GDP,
THUS ALLOWING THE REGENERATION OF EF-1-ALPHA/GTP THAT COULD THEN
BE USED TO FORM THE TERNARY COMPLEX EF-1-ALPHA/GTP/AA/TRNA.
CC -1- SUBUNIT: HOMODIMER.
CC -1- MASS SPECTROMETRY: MW=10006; METHOD=Electrospray.
CC -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL: X76768; CAA54164.1; -.
DR InterPro: IPR001326; EF1BD.
DR Pfam: PF00736; EF1BD; 1.
KW Elongation factor; Protein biosynthesis; Complete proteome.
FT INIT_MET 0
FT SEQUENCE 90 AA; 10005 MW; 24CB357FD3259C8A CRC64;

Query Match 100.0%; Score 20; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 UNLD 4
DB 16 UNLD 19

RESULT 9
EF1B_PYRAB STANDARD; PRT; 91 AA.
ID EF1B_PYRAB
AC O9V2P6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor 1-beta (EF-1-beta) (aEF-1beta).
GN EF1B OR PAB3009.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Pyrococcus.
OX NCBI_TaxID=22992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Helling R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROMOTES THE EXCHANGE OF GDP FOR GTP IN EF-1-ALPHA/GDP,
THUS ALLOWING THE REGENERATION OF EF-1-ALPHA/GTP THAT COULD THEN
BE USED TO FORM THE TERNARY COMPLEX EF-1-ALPHA/GTP/AA/TRNA (BY
SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL: AJ248283; CAB48952.1; -.
DR InterPro: IPR001326; EF1BD.
DR Pfam: PF00736; EF1BD; 1.
KW Elongation factor; Protein biosynthesis; Complete proteome.
SQ SEQUENCE 91 AA; 10239 MW; 45A032D81D54F1EA CRC64;

Query Match 100.0%; Score 20; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 UNLD 4
DB 19 UNLD 22

```

```

RESULT 10
RL15_STRSC STANDARD; PRT; 96 AA.
ID RL15_STRSC
AC P43415;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 50S ribosomal protein L15 (fragment).
GN rplO.
OS Streptomyces scabies.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae.
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1930;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FL1;
RX MEDLINE=96001249; PubMed=7557484;
RA Hale V.A., O'Brien I., Schotet J.L.;
RT "Cloning and sequencing of a secY homolog from Streptomyces scabies.";
RL Gene 163:87-92(1995).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19606; AA85556.1; -
DR InterPro: IPR001196; Ribosomal_L15.
DR Pfam: PF00256; L15; 1.
DR Prosite: PS00475; RIBOSOMAL_L15; 1.
KW Ribosomal protein; rRNA-binding.
FT NON_TER
SQ SEQUENCE 96 AA; 10075 MW; 0BAAC499331BF70E CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 96;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 UNLD 4
DB 26 UNLD 29

RESULT 11
SR19_ARCFU STANDARD; PRT; 104 AA.
ID SR19_ARCFU
AC O29010;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Signal recognition particle 19 kDa protein (SRP19).
GN SRP19 OR AFI258.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E., Kirschum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D., Fleischmann R.D., Kerlavage A.R., Graham D.E., Kyprides N.C., Kikness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B., Peterson S., Reich C.I., McNeill L.K., Badger J.H., Glodok A., Zhou L.,

```

```

RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=20150251; PubMed=10684931;
RA Bhuiyan S.H., Gowda K., Hotokezaka H., Zwiab C.;
RT "Assembly of archaeal signal recognition particle from recombinant
RT components.";
RL Nucleic Acids Res. 28:1365-1373(2000).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=20496765; PubMed=11041851;
RA Diener J.L., Wilson C.;
RT "Role of SRP19 in assembly of the Archaeoglobus fulgidus signal
RT recognition particle.";
RL Biochemistry 39:12862-12874(2000).
CC -1- FUNCTION: SIGNAL-RECOGNITION-PARTICLE ASSEMBLY, BINDS DIRECTLY
CC TO 7S RNA AND MEDIATES BINDING OF THE 54 KDA SUBUNIT OF THE SRP.
CC -1- SUBUNIT: ARCHAEOAL SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA
CC MOLECULE OF 300 NUCLEOTIDES AND TWO PROTEIN SUBUNITS. SRP54 AND
CC SRP19.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE SRP19 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001017; AAB89988.1; -
DR TIGR: AF1258; -
DR InterPro: IPR002778; SRP19.
DR Pfam: PF01922; SRP19; 1.
DR Prodom: PD006609; SRP19; 1.
KW Signal recognition particle; RNA-binding; Ribonucleoprotein; Complete proteome.
SQ SEQUENCE 104 AA; 12405 MW; 72D5DDAB84E89E64 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 104;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 UNLD 4
DB 9 UNLD 12

RESULT 12
HAL2_CANAL STANDARD; PRT; 105 AA.
ID HAL2_CANAL
AC P46594;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Halotolerance protein HAL2 (fragment).
GN HAL2 OR MET22.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WO-1;
RX MEDLINE=95242840; PubMed=7725800;

```

RA Sychorova H., Souciet J.-L.;
 RT "CAN1, a gene encoding a permease for basic amino acids in Candida
 RT albicans.";
 RL Yeast 10:1647-1651(1994).
 CC -1- FUNCTION: INVOLVED IN SALT TOLERANCE AND IN METHIONINE
 CC BIOSYNTHESIS (BY SIMILARITY).
 CC -1- PATHWAY: METHIONINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE INOSITOL MONOPHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X76689; -; NOT_ANNOTATED_CDS.
 DR HSSP: P29218; 1MB.
 DR InterPro: IPR000760; Inositol_P.
 DR Pfam: PF00459; Inositol_P; 1.
 DR PROSITE: PS00629; IMP_1; PARTIAL.
 DR PROSITE: PS00630; IMP_2; 1.
 KW Methionine biosynthesis.
 FT NON_TER
 SO SEQUENCE 105 AA; 11338 MW; EB6938DD4F262F1 CRC64;
 Query Match 100.0%; Score 20; DB 1; Length 105;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 UNLD 4
 Db 9 UNLD 12
 Db 9 UNLD 12
 RESULT 13
 ELIE_BACSU STANDARD; PRT; 106 AA.
 ID ELIE_BACSU
 AC P24502;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Flagellar hook-basal body complex protein flie.
 GN FLIE.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91285431; PubMed-1905667;
 RA Zuberi A.R., Ying C., Bischoff D.S., Ordal G.W.;
 RT "Gene-protein relationships in the flagellar hook-basal body complex
 RT of Bacillus subtilis: sequences of the fligB, fligC, fligE, flie and
 RT flif genes.";
 RL Gene 101:23-31(1991).
 CC -1- SIMILARITY: BELONGS TO THE FLIE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M54965; AAA22443.1; -;
 DR EMBL: Z99112; CAB13493.1; -;
 DR PIR: JG0021; JG0021.
 DR Subtilisin; BG10239; flie.
 DR InterPro: IPR001624; flie.
 DR Pfam: PF02049; flie; 1.

DR PRINTS: PR01006; FLGHOOKFLIE.
 KW Flagella: Complete proteome.
 SO SEQUENCE 106 AA; 11577 MW; 7A3F1B7FE653AC9C CRC64;
 Query Match 100.0%; Score 20; DB 1; Length 106;
 Best Local Similarity 100.0%; Pred. No. 93;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 UNLD 4
 Db 68 UNLD 71
 Db 68 UNLD 71
 RESULT 14
 SPEH_ARCFU STANDARD; PRT; 112 AA.
 ID SPEH_ARCFU
 AC 028663;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50) (AdometDC)
 DE (SampC) [Constraints: S-adenosylmethionine decarboxylase beta chain; S-
 DE adenosylmethionine decarboxylase alpha chain].
 GN SPEH OR AF1610.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OX NCBI_TaxID=2234;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 MEDLINE-98049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Kirkness E.F., Dougeny B.A., McNeill L.K., Badger J.H., Glöckner A., Zhou L.,
 RA Peterson S., Reich C.I., McNeil L.K., Weidman J.F., McDonald L., Ullrich T.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Ullrich T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujil C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -1- FUNCTION: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE
 CC AMINO GROUP. MOIETY REQUIRED FOR SPERMIDINE BIOSYNTHESIS FROM
 CC PUTRESCINE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine = (5'-deoxy-5'-
 CC adenosyl)(3'-aminopropyl) methylsulfonium salt + CO(2).
 CC -1- COFACTOR: Pyruvoyl group (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC ADOMETDC FAMILY. SUBFAMILY
 CC 1
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000991; AAB89640.1; -;
 DR TIGR: AF1610; -;
 DR InterPro: IPR003826; DUF206.
 KW Spermidine biosynthesis; Lyase; Decarboxylase; Lysogen; Pyruvate;
 KW Complete proteome.
 FT CHAIN 1 61 S-ADENOSYLMETHIONINE DECARBOXYLASE BETA
 FT CHAIN 62 112 CHAIN (BY SIMILARITY).
 FT FT S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA
 FT SITE 61 62 CHAIN (BY SIMILARITY).
 FT SITE 61 62 CLEAVAGE (NONHYDROLYTIC) (BY SIMILARITY).

FT MOD_RES 62 62 CONVERTED TO A PYRUVOYL GROUP (BY
FT SEQUENCE 112 AA; 12733 MW; 95B250A1306D2F30 CRC64;
SIMILARITY).

Query Match 100.0%; Score 20; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. NO. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
|||||
Db 75 VNLD 78

Db 40 VNLD 43

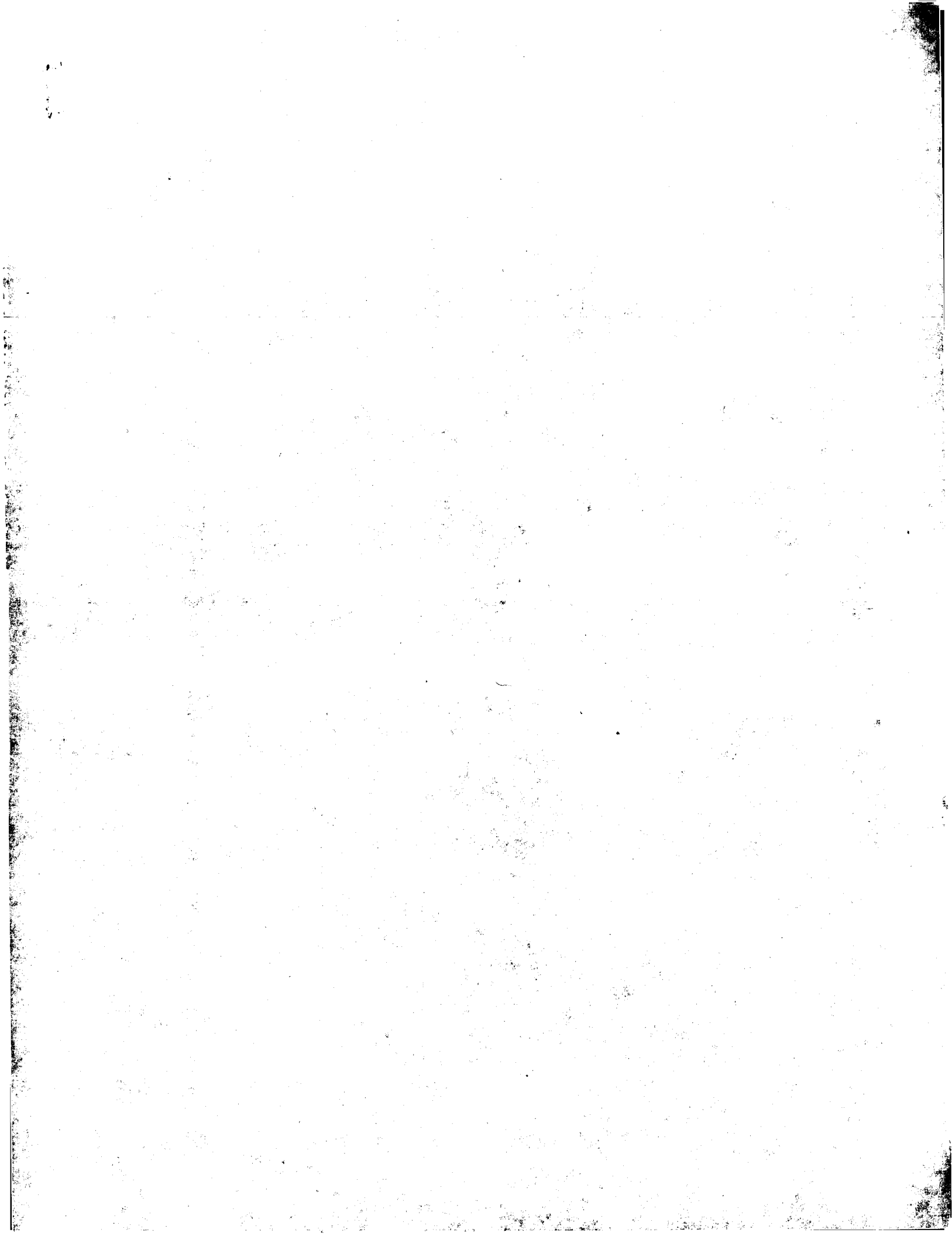
Search completed: October 30, 2002, 12:27:57
Job time : 1.34398 secs

```

RESULT 15
NGF_NAJAT
ID NGF_NAJAT STANDARD; PRT; 116 AA.
AC P21377;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Nerve growth factor (NGF).
OS Naja atra (Chinese cobra), and
OS Naja naja kaouthia (Monocled cobra) (Naja naja siamensis).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX MCB1_TaxID=8636, 8649;
RN [1]
RP SEQUENCE.
RC SPECIES=N.n.kaouthia; TISSUE=Venom;
RX MEDLINE=90147847; PubMed=2619756;
RA Oda T., Ohta M., Inoue S., Ikeda K., Furukawa S., Hayashi K.;
RT "Amino acid sequence of nerve growth factor purified from the venom
of the Formosan cobra Naja naja atra.";
RL Biochem. Int. 19:909-917(1989).
RN [2]
RP SEQUENCE.
RC SPECIES=N.n.kaouthia; TISSUE=Venom;
RX MEDLINE=91138755; PubMed=1995338;
RA Inoue S., Oda T., Koyama J., Ikeda K., Hayashi K.;
RT "Amino acid sequences of nerve growth factors derived from cobra
venoms.";
RL FEBS Lett. 279:38-40(1991).
CC -!- FUNCTION: NERVE GROWTH FACTOR IS IMPORTANT FOR THE DEVELOPMENT AND
CC MAINTENANCE OF THE SYMPATHETIC AND SENSORY NERVOUS SYSTEMS. IT
CC STIMULATES DIVISION AND DIFFERENTIATION OF SYMPATHETIC AND
CC EMBRYONIC SENSOR NEURONS AS WELL AS BASAL FOREBRAIN CHOLINERGIC
CC NEURONS IN THE BRAIN.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SIMILARITY: BELONGS TO THE NGF-BETA FAMILY.
CC PIR: S13965, S13965.
DR HSP: P01139; IBBT.
DR InterPro: IPR002400; GF_cysknot.
DR InterPro: IPR002072; NGF.
DR Pfam: PF00243; NGF_1.
DR PRINTS: PRO0438; GFCYSKNOT.
DR PRINTS: PRO0268; NGF.
DR PRODOM: PRO02052; NGF; 1.
DR SMART: SM00140; NGF; 1.
DR PROSITE: PS00248; NGF_1; 1.
DR PROSITE: PS50270; NGF_2; 1.
KW Growth factor.
FT DISULFID 14 78 BY SIMILARITY.
FT DISULFID 56 106 BY SIMILARITY.
FT DISULFID 66 108 BY SIMILARITY.
SQ SEQUENCE 116 AA; 13064 MW; DAB35421093F3B06 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. NO. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VNLD 4
|||||

```



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:30 ; Search time 1.24816 Seconds

(Without alignments)
554,401 Million cell updates/sec

Title: US-09-724-571-104

Perfect score: 20

Sequence: 1 VNLD 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	32	2 09R507	09R507 aeromonas h
2	20	100.0	33	12 089871	089871 budgerigar
3	20	100.0	33	12 089872	089872 budgerigar
4	20	100.0	33	12 089873	089873 budgerigar
5	20	100.0	33	12 089874	089874 budgerigar
6	20	100.0	33	12 089875	089875 budgerigar
7	20	100.0	33	12 089876	089876 budgerigar
8	20	100.0	33	12 089877	089877 budgerigar
9	20	100.0	33	12 089878	089878 budgerigar
10	20	100.0	33	12 089879	089879 budgerigar
11	20	100.0	33	12 089880	089880 budgerigar
12	20	100.0	33	12 089881	089881 budgerigar
13	20	100.0	33	12 089882	089882 budgerigar
14	20	100.0	33	12 089883	089883 budgerigar
15	20	100.0	33	12 089884	089884 budgerigar
16	20	100.0	33	12 089885	089885 budgerigar

17	20	100.0	33	12 089886	089886 budgerigar
18	20	100.0	33	12 089887	089887 budgerigar
19	20	100.0	33	16 09PH79	09PH79 xylella fas
20	20	100.0	33	16 09RJ3	09RJ3 streptococ
21	20	100.0	40	2 068327	068327 caudobacter
22	20	100.0	41	2 09R2X2	09R2X2 borrelia bu
23	20	100.0	44	2 09S0L7	09S0L7 streptomyc
24	20	100.0	45	4 09HCW4	09HCW4 homo sapien
25	20	100.0	49	5 09VGM9	09VGM9 drosophila
26	20	100.0	49	7 09G1Q2	09G1Q2 eschrichtiu
27	20	100.0	49	8 020180	020180 chlorella v
28	20	100.0	49	13 09PVR4	09PVR4 gallus gall
29	20	100.0	51	2 051106	051106 neisseria m
30	20	100.0	61	2 056859	056859 yersinia en
31	20	100.0	62	10 039072	039072 arbidopsi
32	20	100.0	67	6 09TU51	09TU51 sus scrofa
33	20	100.0	67	13 09IB58	09IB58 xenopus lae
34	20	100.0	67	13 09IB55	09IB55 xenopus lae
35	20	100.0	76	2 09AH64	09AH64 neisseria g
36	20	100.0	79	2 096055	096055 salmonella
37	20	100.0	79	12 084524	084524 paramecium
38	20	100.0	80	16 099SP0	099SP0 staphylococ
39	20	100.0	82	7 09XR70	09XR70 rattus fusc
40	20	100.0	82	7 09XR7	09XR7 rattus fusc
41	20	100.0	82	15 090069	090069 human immun
42	20	100.0	83	2 09ZHA3	09ZHA3 morganella
43	20	100.0	83	2 09XA75	09XA75 streptomyc
44	20	100.0	84	13 09W7I7	09W7I7 gallus gall
45	20	100.0	84	16 053502	053502 mycobacteri

ALIGNMENTS

RESULT 1

ID 09R507 PRELIMINARY; PRT; 32 AA.
AC 09R507;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE S-LAYER PROTEIN.
OS Aeromonas hydrophila.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
OC Aeromonas.
OX NCBI_TaxID=644;
RN [1]
RP SEQUENCE.
RX MEDLINE=92104988; PubMed=1370287;
RA Kostzyńska M., Docley J.S., Shimojo T., Sakata T., Trust T.J.;
RT "Antigenic diversity of the S-layer proteins from pathogenic strains
RT of Aeromonas hydrophila and Aeromonas veronii biotype sobria.";
RL J. Bacteriol. 174:40-47(1992).
SQ SEQUENCE 32 AA; 2931 MW; C1E852A6DA2C8A5 CRC64;

Query Match 100.0%; Score 20; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
Db 1 VNLD 4

RESULT 2

ID 089871 PRELIMINARY; PRT; 33 AA.
AC 089871;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).

```

OC VIRUSES; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=MCFL97;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF054369; AAC33609.1; -.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFAA79 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
Db 27 VNLD 30

RESULT 3
089872 PRELIMINARY; PRT; 33 AA.
AC 089872;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgetigiar fledgling disease virus (BFDV).
OC VIRUSES; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=LBMI92;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF054370; AAC33610.1; -.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFAA79 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
Db 27 VNLD 30

RESULT 4
089873 PRELIMINARY; PRT; 33 AA.
AC 089873;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgetigiar fledgling disease virus (BFDV).
OC VIRUSES; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=GC4292;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF054371; AAC33611.1; -.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFAA79 CRC64;

```

```

SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFAA79 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
Db 27 VNLD 30

RESULT 5
089874 PRELIMINARY; PRT; 33 AA.
AC 089874;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgetigiar fledgling disease virus (BFDV).
OC VIRUSES; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=ECTX91;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF054372; AAC33612.1; -.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFAA79 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
Db 27 VNLD 30

RESULT 6
089875 PRELIMINARY; PRT; 33 AA.
AC 089875;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgetigiar fledgling disease virus (BFDV).
OC VIRUSES; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=LB85;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF054373; AAC33613.1; -.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFAA79 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
Db 27 VNLD 30

```

```
RESULT 7
089876 ID 089876 PRELIMINARY; PRT; 33 AA.
AC 089876;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=XCFL87;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054374; AAC33614.1; -.
FT NON_TER 1 1
FT 33 33
SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFAA79 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
Db 27 VNLD 30

RESULT 8
089877 ID 089877 PRELIMINARY; PRT; 33 AA.
AC 089877;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCIL88;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054375; AAC33615.1; -.
FT NON_TER 1 1
FT 33 33
SQ SEQUENCE 33 AA; 3697 MW; 2034126F3BBFA88 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
Db 27 VNLD 30

RESULT 9
089878 ID 089878 PRELIMINARY; PRT; 33 AA.
AC 089878;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDTX88;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054376; AAC33616.1; -.
FT NON_TER 1 1
FT 33 33
SQ SEQUENCE 33 AA; 3697 MW; 2034126F3BBFA88 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
Db 27 VNLD 30

RESULT 10
089879 ID 089879 PRELIMINARY; PRT; 33 AA.
AC 089879;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RNA191, LRTX93;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054377; AAC33617.1; -.
FT NON_TER 1 1
FT 33 33
SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFAA79 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNLD 4
Db 27 VNLD 30

RESULT 11
089880 ID 089880 PRELIMINARY; PRT; 33 AA.
AC 089880;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases)
DE LARGE T ANTIGEN (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDTX89;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054378; AAC33618.1; -.
FT NON_TER 1 1
FT 33 33
SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFAA79 CRC64;
```

Query Match 100.0%; Score 20; DB 12; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
 ||||
 DB 27 VNLD 30

RESULT 12
 ID 089881 PRELIMINARY; PRT; 33 AA.
 AC 089881;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LARGE T ANTIGEN (FRAGMENT).
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 NCBI_TaxID=10625;
 RN NCBI_TaxID=10625;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BDGA81-A;
 RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
 RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus."
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF054379; AAC33619.1; -.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFA79 CRC64;

Query Match 100.0%; Score 20; DB 12; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
 ||||
 DB 27 VNLD 30

RESULT 13
 ID 089882 PRELIMINARY; PRT; 33 AA.
 AC 089882;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LARGE T ANTIGEN (FRAGMENT).
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 NCBI_TaxID=10625;
 RN NCBI_TaxID=10625;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BDGA81-B;
 RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
 RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus."
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF054380; AAC33620.1; -.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFA79 CRC64;

Query Match 100.0%; Score 20; DB 12; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
 ||||
 DB 27 VNLD 30

RESULT 14
 C93883

ID 089883 PRELIMINARY; PRT; 33 AA.

AC 089883;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LARGE T ANTIGEN (FRAGMENT).
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 NCBI_TaxID=10625;
 RN NCBI_TaxID=10625;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BCFL92;
 RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
 RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus."
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF054381; AAC33621.1; -.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFA79 CRC64;

Query Match 100.0%; Score 20; DB 12; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLD 4
 ||||
 DB 27 VNLD 30

RESULT 15
 ID 089884 PRELIMINARY; PRT; 33 AA.
 AC 089884;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LARGE T ANTIGEN (FRAGMENT).
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 NCBI_TaxID=10625;
 RN NCBI_TaxID=10625;
 RP SEQUENCE FROM N.A.
 RC STRAIN=ECFL91;
 RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
 RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus."
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF054382; AAC33622.1; -.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 3713 MW; ABA4126F3BBFA79 CRC64;

QY 1 VNLD 4
 ||||
 DB 27 VNLD 30

Search completed: October 30, 2002, 12:30:17
 Job time: 3.24816 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:24:20 ; Search time 51.4447 Seconds

(without alignments)
171,400 Million cell updates/sec

Title: US-09-724-571-75

Perfect score: 1907
1 MVDNLKRGSGQGYVEMTVG.....VVFDRKRKRIGFAVSACHVH 361

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6C_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1901	99.7	501	US-09-009-191-2	Sequence 2, Appl1
2	1890	99.1	774	US-09-009-191-4	Sequence 4, Appl1
3	1106	58.0	518	US-08-999-723-2	Sequence 2, Appl1
4	1106	58.0	518	US-09-434-427-2	Sequence 2, Appl1
5	1088	57.1	514	US-09-717-432-2	Sequence 2, Appl1
6	1088	57.1	514	US-09-912-484-2	Sequence 2, Appl1
7	298.5	15.7	396	US-08-208-007A-13	Sequence 13, Appl1
8	298.5	15.7	396	US-09-032-523-9	Sequence 9, Appl1
9	295.5	15.5	412	US-08-208-007A-12	Sequence 12, Appl1
10	295.5	15.5	412	US-08-974-691-4	Sequence 4, Appl1
11	279.5	14.7	458	US-09-640-305-6	Sequence 6, Appl1
12	273.5	14.3	409	US-08-360-673-6	Sequence 8, Appl1
13	273.5	14.3	409	US-08-846-021A-8	Sequence 8, Appl1
14	273.5	14.3	427	US-08-088-633-2	Sequence 2, Appl1
15	270	14.2	410	US-08-088-633-2	Sequence 2, Appl1
16	270	14.2	410	US-08-245-756-2	Sequence 2, Appl1
17	270	14.2	410	US-08-441-750-2	Sequence 2, Appl1
18	270	14.2	410	US-08-441-751-2	Sequence 2, Appl1
19	270	14.2	410	PCT-US92-02521-2	Sequence 2, Appl1
20	249	13.1	349	US-09-032-523-3	Sequence 3, Appl1
21	227	11.9	398	US-08-328-314-2	Sequence 2, Appl1
22	227	11.9	398	US-08-731-045-2	Sequence 2, Appl1
23	212	11.1	419	US-08-974-691-3	Sequence 3, Appl1
24	210	11.0	397	US-09-079-415-2	Sequence 2, Appl1
25	195.5	10.3	427	US-07-958-222A-2	Sequence 2, Appl1
26	194.5	10.2	430	US-08-535-237-2	Sequence 2, Appl1
27	193	10.1	330	US-08-115-753-1	Sequence 1, Appl1

ALIGNMENTS

28	193	10.1	419	3	US-08-115-753-2	Sequence 2, Appl1
29	193	10.1	419	3	US-08-115-753-33	Sequence 33, Appl1
30	184	9.6	445	4	US-08-974-691-6	Sequence 6, Appl1
31	184	9.6	451	4	US-08-974-691-2	Sequence 2, Appl1
32	180	9.4	420	4	US-09-008-227A-4	Sequence 4, Appl1
33	180	9.4	420	4	US-08-974-691-8	Sequence 8, Appl1
34	177.5	9.3	395	1	US-08-723-938-3	Sequence 3, Appl1
35	177.5	9.3	395	1	US-09-080-538-3	Sequence 3, Appl1
36	149	7.8	437	4	US-09-353-332-2	Sequence 2, Appl1
37	129.5	6.8	140	3	US-09-211-631-13	Sequence 13, Appl1
38	129.5	6.8	140	4	US-09-265-628-13	Sequence 13, Appl1
39	129.5	6.8	140	4	US-09-532-803-6	Sequence 11, Appl1
40	129.5	6.8	140	4	US-09-532-803-6	Sequence 6, Appl1
41	129.5	6.8	140	4	US-09-653-403-14	Sequence 14, Appl1
42	97	5.1	1030	4	US-09-091-117-2	Sequence 2, Appl1
43	95.5	5.0	280	4	US-09-160-246-14	Sequence 14, Appl1
44	85	4.5	377	2	US-08-853-659A-41	Sequence 41, Appl1
45	84.5	4.4	1097	2	US-08-680-326-39	Sequence 39, Appl1

RESULT 1

US-09-009-191-2

Sequence 2, Application US/09009191

Patent No. 6319689

GENERAL INFORMATION:

APPLICANT: POWELL, DAVID

APPLICANT: CHAPMAN, CONRAD

APPLICANT: MURPHY, KAY

APPLICANT: SMITH, TRUDI

TITLE OF INVENTION: ASP2

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESS: RATNER & PRESTIA

STREET: P.O. BOX 980

CITY: VALLEY FORGE

STATE: PA

COUNTRY: USA

ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/009,191

FILING DATE: 20-JAN-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: UK 9701684.4

FILING DATE: 128-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F

REGISTRATION NUMBER: 23,031

REFERENCE/DOCKET NUMBER: GH-70368

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-407-0700

TELEFAX: 610-407-0701

TELEX: 846169

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 501 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-009-191-2

Query Match

Best Local Similarity 99.7%; Score 1901; DB 4;

Matches 360; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 MVDNLKSGSGGYVEMTVGSPQTLNLLVDTGSSNFVAGAAPHPFLHRYQROLSTYR 60
Db      63 MVDNLKSGSGGYVEMTVGSPQTLNLLVDTGSSNFVAGAAPHPFLHRYQROLSTYR 122
QY      61 DLKRGVYVPYTGKWEGLGELTDLVSIIPHGPNTVVRANIAATIESDKFFINGSNMEGILGL 120
Db      123 DLKRGVYVPYTGKWEGLGELTDLVSIIPHGPNTVVRANIAATIESDKFFINGSNMEGILGL 182
QY      121 AYAETARPDDSLPEFPDLSLVKQTHVPLNLSLQCGAGFPLNOSVILASVGSMTIGIDH 180
Db      183 AYAETARPDDSLPEFPDLSLVKQTHVPLNLSLQCGAGFPLNOSVILASVGSMTIGIDH 242
QY      181 SLVYGLMYPPIRREMYEVIIVRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKKYF 240
Db      243 SLVYGLMYPPIRREMYEVIIVRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKKYF 302
QY      241 EAAVKSIAASSTKFKPDGFWLGEOLVCMQAGTTPMNIFFVISLYLMGEVYNOSFRITIL 300
Db      303 EAAVKSIAASSTKFKPDGFWLGEOLVCMQAGTTPMNIFFVISLYLMGEVYNOSFRITIL 362
QY      301 PQOYLRPVEDVATSDODCKKFAISOSSTGTVMGAVIMEGFYVFDRAKRRIIGFAVSACHV 360
Db      363 PQOYLRPVEDVATSDODCKKFAISOSSTGTVMGAVIMEGFYVFDRAKRRIIGFAVSACHV 422
QY      361 H 361
Db      423 H 423

```

```

RESULT 2
US-09-009-191-4
; Sequence 4, Application US/09009191
; Patent No. 6319689

```

```

GENERAL INFORMATION:
APPLICANT: POWELL, DAVID
APPLICANT: CHAPMAN, CONRAD
APPLICANT: MURPHY, KAY
APPLICANT: SMITH, TRUDI
TITLE OF INVENTION: ASP2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,191
FILING DATE: 20-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 9701684.4
FILING DATE: 28-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70368
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 amino acids
TYPE: amino acid
STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-009-191-4

```

```

Query Match          99.1%; Score 1890; DB 4; Length 774;
Best Local Similarity 99.2%; Pred. No. 3,7e-186;
Matches 358; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 MVDNLKSGSGGYVEMTVGSPQTLNLLVDTGSSNFVAGAAPHPFLHRYQROLSTYR 60
Db      6 MVDNLKSGSGGYVEMTVGSPQTLNLLVDTGSSNFVAGAAPHPFLHRYQROLSTYR 65
QY      61 DLKRGVYVPYTGKWEGLGELTDLVSIIPHGPNTVVRANIAATIESDKFFINGSNMEGILGL 120
Db      126 AYAETARPDDSLPEFPDLSLVKQTHVPLNLSLQCGAGFPLNOSVILASVGSMTIGIDH 185
QY      121 AYAETARPDDSLPEFPDLSLVKQTHVPLNLSLQCGAGFPLNOSVILASVGSMTIGIDH 180
Db      66 DLKRGVYVPYTGKWEGLGELTDLVSIIPHGPNTVVRANIAATIESDKFFINGSNMEGILGL 125
QY      181 SLVYGLMYPPIRREMYEVIIVRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKKYF 240
Db      186 SLVYGLMYPPIRREMYEVIIVRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKKYF 245
QY      241 EAAVKSIAASSTKFKPDGFWLGEOLVCMQAGTTPMNIFFVISLYLMGEVYNOSFRITIL 300
Db      246 EAAVKSIAASSTKFKPDGFWLGEOLVCMQAGTTPMNIFFVISLYLMGEVYNOSFRITIL 305
QY      301 PQOYLRPVEDVATSDODCKKFAISOSSTGTVMGAVIMEGFYVFDRAKRRIIGFAVSACHV 360
Db      306 PQOYLRPVEDVATSDODCKKFAISOSSTGTVMGAVIMEGFYVFDRAKRRIIGFAVSACHV 365
QY      361 H 361
Db      366 H 366

```

```

RESULT 3
US-08-999-723-2
; Sequence 2, Application US/08999723A
; Patent No. 6025180

```

```

GENERAL INFORMATION:
APPLICANT: Powell, David J.
APPLICANT: Southan, Christopher
APPLICANT: Chapman, Conrad G.
APPLICANT: Pyans, Joanne R.
TITLE OF INVENTION: ASP1
FILE REFERENCE: GH/0262
CURRENT APPLICATION NUMBER: US/08/999,723A
CURRENT FILING DATE: 1997-10-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 518
TYPE: PRT
ORGANISM: Homo sapiens
US-08-999-723-2

```

```

Query Match          58.0%; Score 1106; DB 3; Length 518;
Best Local Similarity 56.4%; Pred. No. 1.5e-105;
Matches 202; Conservative 60; Mismatches 92; Indels 4; Gaps 2;

```

```

QY      1 MVDNLKSGSGGYVEMTVGSPQTLNLLVDTGSSNFVAGAAPHPFLHRYQROLSTYR 60
Db      80 MVDNLGDSGRGYLLEMLIGTPQKLOLLVDTGSSNFVAGATPHSYITDYFTERSSTYR 139
QY      61 DLKRGVYVPYTGKWEGLGELTDLVSIIPHGPNTVVRANIAATIESDKFFINGSNMEGILGL 120
Db      140 SKGPDVYVYKTYGSGWTFGVGEDLVITPKGNTSFLVIAITIFFSENFLPGIKWNGILGL 199
QY      121 AYAETARPDDSLPEFPDLSLVKQTHVPLNLSLQCGAGFPLNOSVILASVGSMTIGIDH 180
Db      200 AYAETARPDDSLPEFPDLSLVKQTHVPLNLSLQCGAGFPLNOSVILASVGSMTIGIDH 256

```


Best Local Similarity 55.9%; Pred. No. 1,1e-103;
Matches 200; Conservative 60; Mismatches 94; Indels 4; Gaps 2;

QY 1 MYDNRKSGGGYVEMVSGPQTLNLVDTGSSNFVAGAPHEFLHRRYQROLSSYR 60
Db 76 MYDNRKSGGGYVEMVSGPQTLNLVDTGSSNFVAGAPHEFLHRRYQROLSSYR 135
QY 61 DLKRGVYVYOGKMEGLGFDLVSIPLGPNVTVRANIAITSDKFFINGSNMWGL 120
Db 136 SKGPDVYKYTGQSGMTGVGDLVYIPKGFNSFLVNIATIFESNFFLPGLKMGIL 195
QY 121 AYAEIARDDSLPEFDSLVKQTHVNLFSLOLCGAGFPLNOSVLAAGSMITIGIDH 180
Db 196 AYAAIAKSSSLPEFDSLVKQTHVNLFSLOLCGAGFPLNOSVLAAGSMITIGIDH 252
QY 181 SLTYSGLWTPRRWYEVILVIRREINGODLKMCKEYNTDKSIVDSGTTLRLPKYF 240
Db 253 SLTYSGLWTPRRWYEVILVIRREINGODLKMCKEYNTDKSIVDSGTTLRLPKYF 312
QY 241 EAAVSIKAASTKPEPDGFMGLGQVLCWQAGTTPMNIIPVLSLYLMGEVTLNQSFRITL 300
Db 313 DAVVAIVARTSLIPEFSDGFMGLGQVLCWQAGTTPMNIIPVLSLYLMGEVTLNQSFRITL 372
QY 301 PQQYLRPEVDVATSDCCYKFAISQSGTGVGAVIMSGFVYFPRARRKIFAVSAC 358
Db 373 POLYIOPMAGAFNY-ECYRFGISSSTNALVIGATVMEGFYVFDRAORVGFVAVSPC 429

RESULT 7

US-08-208-007A-13
Sequence 13, Application US/08208007A
Patent No. 5501969

GENERAL INFORMATION:
APPLICANT: HASTINGS, ET AL.

TITLE OF INVENTION: Human Osteoclast-Derived Cathepsin

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/208,007A

FILING DATE: March 8, 1994

CLASSIFICATION: A35

PRIOR APPLICATION DATA:
APPLICATION NUMBER: No. 5501969e

FILING DATE: No. 5501969e

ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-95

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 396 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN

US-08-208-007A-13

Query Match 15.7%; Score 298.5; DB 1; Length 396;

Best Local Similarity 26.7%; Pred. No. 2e-22;
Matches 96; Conservative 61; Mismatches 141; Indels 61; Gaps 14;

QY 13 YVEWTVGSPQTLNLVDTGSSNFVAGAPHEFLHRRYQROLSSYRDLRKGYV 68
Db 78 YVEWTVGSPQTLNLVDTGSSNFVAGAPHEFLHRRYQROLSSYRDLRKGYV 137
QY 69 PYTGKMEGLGFDLVSIPLGPNVTVRANIAITSDKFFINGSNMWGLAELIAR 128
Db 138 PYTGKMEGLGFDLVSIPLGPNVTVRANIAITSDKFFINGSNMWGLAELIAR 193
QY 129 DDLPEFDSLVKQTHVNLFSLOLCGAGFPLNOSVLAAGSMITIGIDHSLYTSGLW 188
Db 194 DDLPEFDSLVKQTHVNLFSLOLCGAGFPLNOSVLAAGSMITIGIDHSLYTSGLW 246
QY 189 YTPIRREWYEVILVIRREINGODLKMCKEYNTDKSIVDSGTTLRLPKYF 248
Db 247 YTPIRREWYEVILVIRREINGODLKMCKEYNTDKSIVDSGTTLRLPKYF 302
QY 249 AASTKPEPDGFMGLGQVLCWQAGTTPMNIIPVLSLYLMGEVTLNQSFRITL 308
Db 303 AASTKPEPDGFMGLGQVLCWQAGTTPMNIIPVLSLYLMGEVTLNQSFRITL 341
QY 309 EDVATSDCCYKFAISQSGTGVGAVIMSGFVYFPRARRKIFAVSAC 354
Db 342 EDVATSDCCYKFAISQSGTGVGAVIMSGFVYFPRARRKIFAVSAC 392

RESULT 8

US-09-032-523-9
Sequence 9, Application US/09032523
Patent No. 6232454

GENERAL INFORMATION:
APPLICANT: Bandman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Guebler, Karl

TITLE OF INVENTION: HUMAN PROTEINASE MOLECULES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

OPERATING SYSTEM: IBM Compatible

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,523

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0479 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 396 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 181994
US-09-032-523-9

Query Match 15.7% Score 298.5; DB 4; Length 396;
Best Local Similarity 26.7%; Pred. No. 2e-22;
Matches 96; Conservative 61; Mismatches 141; Indels 61; Gaps 14;

QY 13 YVEVETVSPQTLNLIIVDSSNFAYGA----APHPFLRYQROLSTTYRDLRKGVY 68
D 13 YVEVETVSPQTLNLIIVDSSNFAYGA----APHPFLRYQROLSTTYRDLRKGVY 68
D 78 YFGISISGPPNFTVIFDSSNLMVPSYCTSPACKTSRQPSQSTYSOPGGSFSI 137
QY 69 PYTGKMEGELTDLVSIPIHPNVTYRANIAITSDKFFINSNMGEIGLVAETARP 128
D 138 QYGTSLGSIIGADQVAV-EGIFVVGQGFESVTEPGQITVD-AEDDILGLIYPSLA-- 193
QY 129 DDLSEPFDSLVKQTHPNLFSIQLCGAGFPLNOSVLA SVGSMITIGIDHSLYGSIM 188
D 194 VCGTTPFDNMMAO-----NLVDLPMSVYSSNPE--GGAGSELIIFGVDHSHFSGSIN 246
QY 189 YTPIRREMYEVIIVRVEINGODLKMCKEYNDKSIYDSGTNLRPKKVEFAAKSK 248
D 247 WVPYTKQAYQIALDNQVGG--TYMFCSE--GCQAVDTGTSITLPSPDKIKOLONALG 302
QY 249 AASSTEFPPGFMIGBOLVCMQAGTTPWNIPIVSIYLMGEVTVNOSFRITILPOQYLREV 308
D 303 AAP-----VDEGAYVE-----CANLNMVDPVFTING-----VPYTLSPYAV--TL 341
QY 309 EDVATSDODCYKFAISQSTG-----TWGAVIMEGFEVVDRAKRRIGFA 354
D 342 LDFVDGMQFC-----SSGFGGLDIHPAGFLMILGDVFTKQFYVSPDRGNRVGIA 392

RESULT 9

US-08-208-007A-12
Sequence 12, Application US/08208007A
Patent No. 5501969

GENERAL INFORMATION:
APPLICANT: HASTINGS, ET AL.

TITLE OF INVENTION: Human Osteoclast-Derived Cathepsin

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER PARK ROAD
CITY: ROSELAND

STATE: NEW JERSEY
COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/208, 007A

FILING DATE: March 8, 1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: No. 5501969e

FILING DATE: No. 5501969e
ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-95
TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:

LENGTH: 412 AMINO ACIDS
TYPE: AMINO ACID

STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-208-007A-12

Query Match 15.5% Score 295.5; DB 1; Length 412;
Best Local Similarity 28.5%; Pred. No. 4.4e-22;
Matches 105; Conservative 58; Mismatches 146; Indels 59; Gaps 15;

QY 13 YVEVETVSPQTLNLIIVDSSNFAYGAAPHPFL-----HRYQROLSTTYRDLRKGV 66
D 79 YGEIGIGIPPCCFYVVDGTSSNLMVPSICKLIDICWIIHKYNSDKSSTYKNGTSF 138
QY 67 YPYTGKMEGELTDLVSIPIHPNVTYRANIAITSDKFFINSNMGEI 117
D 139 DIHYSGSLSGYLSODFTVSPQSSASALAGVYKVERQGEANTKQGIFFIAKPDGI 198
QY 118 LGLAVAEIARPDLSLEPFDSLVKQTHV-PNLFSLQLCGAGFPLNOSVLA SVGSMITIG 176
D 199 LGMATPRIS--VNNVLPVFDNLMOQKLVQNIFFSYL-----SRDPAQPGGELMLG 248
QY 177 GIDHSLYGSIMVTPPIREMYEVIIVRVEI-NGODLKMCKEYNDKSIYDSGTNLR 235
D 249 GDSKRYKXGSLYLNVTKRAYQVHLDOVEVASGLTL---CKE--GCEALVDIGTSLMG 303
QY 236 PKVFEAAVKSIKAASSTEFPPDGFMLGBOLV-CWQAGTTPWNIPIVSIYLMGEVTVNOS 294
D 304 PVDEYRELQKAIGAVPLIQ-----GEYMIPEKYST-----LPAITLKLGG---KG 346
QY 295 FRITILPOOYLRPEDVATSDODCYKFAISQ-----SSTGVMGAVIMEGFEVYVFDARK 349
D 347 YKLS--PEDYTLKVSQAGKITL--CLSGFMGMDIPPSGFLMILGDVFTKQFYVFDNRDN 402
QY 350 RIGFAVSA 357
D 403 RVGFAEAA 410

RESULT 10

US-08-974-691-4
Sequence 4, Application US/08974691
Patent No. 6225103

GENERAL INFORMATION:
APPLICANT: Keolsch, Gerald

APPLICANT: Lin, Xinli

APPLICANT: Tang, Jordan

TITLE OF INVENTION: Cloning and Characterization of Napsin

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 W. Peachtree

STREET: St.
CITY: Atlanta

STATE: GA
COUNTRY: USA

ZIP: 30309-3450

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,691

FILING DATE: 20-NOV-1997
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,196

FILING DATE: 20-NOV-1996
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/046,126
FILING DATE: 09-MAY-1997

ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.

Best Local Similarity 25.3%; Pred. No. 8.1e-20;
Matches 92; Conservative 59; Mismatches 144; Indels 69; Gaps 11;

QY 13 YVEMTVGSPQTLNLTVDGSSNFAVGAAP-----HPFLHRYQROSLSTYRDLRKGYV 68
Db 96 YFEITILGSPQSFVKYILDGSSNLMWPFAEGSLACFLHRYDHEASTYKANGSEPAI 155
QY 69 PYOGKWEGLGTDLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGILGLAETIAR 128
Db 156 QVSGSLEGVSRDLTLTI--GDVLIPODFAEATSEPLAFAGKFGDILGLADTSIS-V 212
QY 129 DDLSEFFDSLVKQTHVNLFSIQLCGAGFPLNOSSEVLASVGSMTIIGIDHSITGSLW 188
Db 213 NRIYPPVYNAIKNKLDDPVFA-----FYLGDSD-KSEGGGASFGIDEKTYTEIT 264
QY 189 YPIRREMYEYIIVREINQDLKMDCKEYNDKSIYDSTGNLRPKKVFEAAYSIK 248
Db 265 WLPVRKKAWE-----VFEGIGLGEYATLEGHGAALDTGTSIALPSGLAETIAR 319
QY 249 AASSTKFPDGFVLGEOLVCMQAGTTPWNIFPVISLYLMEGVNOSFRIT----- 298
Db 320 AKKG-----WSGQYSVDCESRDS---LPDLTLNFG---YNFTIAYDYTLEVS 363
QY 299 -----ILPQYLRPVEDVATSDQCYKFAISOSTGIVMGAVIMEGYVVEDRARRKRG 353
Db 364 SCISAFTEPMDPEPVGPLA-----IIGDALRKRYSTIYDIGHDAVGL 405
QY 354 AVSA 357
Db 406 AKAA 409

RESULT 13

US-08-360-673-6
Sequence 6, Application US/08360673
Patent No. 5679544

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Yeh, Patrice
TITLE OF INVENTION: MODIFIED KLUYEROMYCES YEASTS, THEIR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360.673
FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR93/00623
FILING DATE: 23-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/07785
FILING DATE: 25-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Smith, Julie K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST92040-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-360-673-6

Query Match 14.3%; Score 273.5; DB 1; Length 409;
Best Local Similarity 25.3%; Pred. No. 8.1e-20;
Matches 92; Conservative 59; Mismatches 144; Indels 69; Gaps 11;

QY 13 YVEMTVGSPQTLNLTVDGSSNFAVGAAP-----HPFLHRYQROSLSTYRDLRKGYV 68
Db 96 YFEITILGSPQSFVKYILDGSSNLMWPFAEGSLACFLHRYDHEASTYKANGSEPAI 155
QY 69 PYOGKWEGLGTDLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGILGLAETIAR 128
Db 156 QVSGSLEGVSRDLTLTI--GDVLIPODFAEATSEPLAFAGKFGDILGLADTSIS-V 212
QY 129 DDLSEFFDSLVKQTHVNLFSIQLCGAGFPLNOSSEVLASVGSMTIIGIDHSITGSLW 188
Db 213 NRIYPPVYNAIKNKLDDPVFA-----FYLGDSD-KSEGGGASFGIDEKTYTEIT 264
QY 189 YPIRREMYEYIIVREINQDLKMDCKEYNDKSIYDSTGNLRPKKVFEAAYSIK 248
Db 265 WLPVRKKAWE-----VFEGIGLGEYATLEGHGAALDTGTSIALPSGLAETIAR 319
QY 249 AASSTKFPDGFVLGEOLVCMQAGTTPWNIFPVISLYLMEGVNOSFRIT----- 298
Db 320 AKKG-----WSGQYSVDCESRDS---LPDLTLNFG---YNFTIAYDYTLEVS 363
QY 299 -----ILPQYLRPVEDVATSDQCYKFAISOSTGIVMGAVIMEGYVVEDRARRKRG 353
Db 364 SCISAFTEPMDPEPVGPLA-----IIGDALRKRYSTIYDIGHDAVGL 405
QY 354 AVSA 357
Db 406 AKAA 409

RESULT 14

US-08-846-021A-8
Sequence 8, Application US/08846021A
Patent No. 5948682

GENERAL INFORMATION:

APPLICANT: Moloney, Maurice M.
TITLE OF INVENTION: Preparation of Heterologous Proteins on
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/846.021A
FILING DATE: April 25, 1997
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Gravelle, Michelle
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 9369-039
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 8:

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:10 ; Search time 31.0442 Seconds

(without alignments)
450.253 Million cell updates/sec

Title: US-09-724-571-75

Perfect score: 1907
Sequence: 1 MVDLNRKSGGGYVEMTVG.....VFDRARRKRGFAVSACHVH 361

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1907	100.0	501	1	BACE_HUMAN
2	1896	99.4	501	1	BACE_MOUSE
3	1890	99.1	501	1	BACE_RAT
4	1106	58.0	518	1	BACE2_HUMAN
5	327	17.1	324	1	PEP1_GADMO
6	306.5	16.1	390	1	CATD_BOVIN
7	306	16.0	367	1	PEP1_RABIT
8	302	15.8	367	1	PEP1_CHICK
9	302	15.8	388	1	PEP4_MACFU
10	298.5	15.7	383	1	PEPE_CHICK
11	298.5	15.7	396	1	CATE_HUMAN
12	296	15.5	387	1	PEP2_RABIT
13	296	15.5	391	1	CATE_CAVPO
14	295.5	15.5	412	1	CATD_HUMAN
15	295	15.5	387	1	PEP4_RABIT
16	287	15.0	407	1	CATD_RAT
17	286	15.0	387	1	PEP3_RABIT
18	285.5	15.0	398	1	CATE_RAT
19	285	14.9	388	1	PEP2_MACFU
20	285	14.9	398	1	CATD_CHICK
21	283	14.8	388	1	PEP4_MACMU
22	281.5	14.8	381	1	CHYM_SHEEP
23	281	14.7	388	1	PEP1_MACFU
24	279.5	14.7	410	1	CATD_MOUSE
25	278.5	14.6	386	1	PEP4_PIG
26	278	14.6	387	1	PEP4_CALJA
27	278	14.6	388	1	PEP4_HUMAN
28	277.5	14.6	396	1	CATE_RABIT
29	277.5	14.6	397	1	CATE_RABIT
30	274	14.4	419	1	CARV_CANAL
31	273.5	14.3	381	1	CHYM_BOVIN
32	273	14.3	388	1	PEP2_RABIT
33	270.5	14.2	377	1	PEP3_MACFU

34	270	14.2	376	1	PAC2_BOVIN
35	267.5	14.0	381	1	CHYM_CALJA
36	266	13.9	405	1	CARP_YEAST
37	264	13.8	396	1	CARP_NEUCR
38	263.5	13.8	388	1	PEPC_CALJA
39	262.5	13.8	394	1	PEPC_CAVPO
40	261.5	13.7	387	1	ASPR_AEDAE
41	261.5	13.7	388	1	PEPC_HUMAN
42	261	13.7	388	1	PAG_HORSE
43	260	13.6	345	1	CATD_PIG
44	259	13.6	365	1	CATD_SHEEP
45	250.5	13.1	496	1	ASPR_ORYSA

ALIGNMENTS

RESULT 1	ID	NAME	STANDARD	PRT	501 AA
AC	P56817	OGUT5			
DT	30-MAY-2000	(Rel. 39, Created)			
DT	30-MAY-2000	(Rel. 39, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)				
DE	(Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl				
DE	protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)				
DE	(Memapsin-2).				
GN	BACE OR BACE1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxId=9606;				
RN	[1]	SEQUENCE FROM N.A. (ISOFORM A).			
RP	TISSUE=Brain;				
RC	MEDLINE=20002972; PubMed=10531052;				
RA	Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,				
RA	Denis P., Teplow D.B., Ross S., Amaratne P., Loeloff R., Luo Y.,				
RA	Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,				
RA	Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,				
RT	Treanor J., Rogers G., Citron M.;				
RT	"Beta-secretase cleavage of Alzheimer's amyloid precursor protein by				
RL	the transmembrane aspartic protease BACE.";				
RL	Science 286:735-741(1999).				
RP	[2]	SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 46-68, AND			
RP	CHARACTERIZATION.				
RC	TISSUE=Brain;				
RC	MEDLINE=20057171; PubMed=10591214;				
RA	Sinha S., Anderson J.P., Barbour R., Basi G.S., Caccavello R.,				
RA	Davis D., Doan M., Dorey H.F., Frigon N., Hong J., Jacobson-Croak K.,				
RA	Jeetun N., Keim P., Knops J., Lieberburg I., Power M., Tan H.,				
RA	Tsutsuno G., Tung J., Schenk D., Seubert P., Stomenssart S.M., Wang S.,				
RA	Walker D., Zhao J., McConlogue L., Varghese J.;				
RT	"Purification and cloning of amyloid precursor protein beta-secretase				
RT	from human brain.";				
RL	Nature 402:537-540(1999).				
RP	[3]	SEQUENCE FROM N.A. (ISOFORM A).			
RP	MEDLINE=20057170; PubMed=10591213;				
RA	van R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,				
RA	Brashers J.R., Stralman N.C., Mathews W.R., Buhl A.E., Carter D.B.,				
RA	Tomasselli A.G., Parodi L.A., Heintz R.L., Gurney M.E.;				
RT	"Membrane anchored aspartyl protease with Alzheimer's disease beta-				
RT	secretase activity.";				
RL	Nature 402:533-537(1999).				
RP	[4]	SEQUENCE FROM N.A. (ISOFORM A).			
RP	MEDLINE=20030166; PubMed=10561122;				
RA	Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,				
RA	Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,				
RA	Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;				

RT "Identification of a novel aspartic proteinase (Asp 2) as beta-secretase.";
 RL Mol. Cell. Neurosci. 14:419-427(1999).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM B).
 RC TISSUE=Brain, and Pancreas;
 RA Michel B., De Pietri Tonelli D., Zaccchetti D., Keller P.;
 RT "New beta-site APP cleaving enzyme isoform (BACE-1B) obtained from human brain and pancreas.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.
 RA MEDLINE-20144060; PubMed-10677483;
 RX Lin X., Koelsch G., Wu S., Downs D., Dashed A., Tang J.;
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
 CC -1- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP, LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: A/BACE-1A (shown here) and B/BACE-1B, are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF190725; AAF04142.1; -
 DR EMBL: AF201468; AAF18982.1; -
 DR EMBL: AF200343; AAF1079.1; -
 DR EMBL: AF204943; AAF26367.1; -
 DR EMBL: AF338816; AAK38374.1; -
 DR EMBL: AF200193; AAF13715.1; -
 DR HSP: P32329; 1YPS.
 DR MIM: 604252; -
 DR InterPro: IPR001969; Asp-protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp; 3.
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; Asp-PROTEASE; 1.
 KW Hydrolyase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
 RL Signal; Alternative splicing.
 FT SIGNAL 1 21
 FT PROPEP 22 45
 FT CHAIN 46 501
 FT DOMAIN 22 457
 FT TRANSMEM 458 478
 FT DOMAIN 479 501
 FT ACT_SITE 93 93
 FT ACT_SITE 289 289
 FT CARBOHYD 153 153
 FT CARBOHYD 172 172
 FT CARBOHYD 223 223
 FT CARBOHYD 354 354
 FT VARSPLIC 190 214
 FT SEQUENCE 501 AA; 55763 MW; 370E4C824ACEF05 CRC64;
 Query Match Score 1907; DB 1; Length 501;
 Best Local Similarity 100.0%; Pred. No. 6e-153;
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 63 MVDNLKSGGQYGVVEMTVGSPQTLNIIIVDTGSSNFVAGADPHPLHRYQRQLSSYTR 122
 Qy 61 DLKRGVYVPTGCKWEGELTDLVSIPIHGPVYVRAIAIAITSDKFFINGSMWEGILGL 120
 Db 123 DLKRGVYVPTGCKWEGELTDLVSIPIHGPVYVRAIAIAITSDKFFINGSMWEGILGL 182
 Qy 121 AYAEIARPDLSLEFFPDLSVKOTHPNLFSLQAGAPFNPOSEVILASVGSMTIGSIDH 180
 Db 183 AYAEIARPDLSLEFFPDLSVKOTHPNLFSLQAGAPFNPOSEVILASVGSMTIGSIDH 242
 Qy 181 SLVYGLMYPPIREMYEYVITVAVETINGODLKMDCKEYVNDKSYVSGTTLRLPKVF 240
 Db 243 SLVYGLMYPPIREMYEYVITVAVETINGODLKMDCKEYVNDKSYVSGTTLRLPKVF 302
 Qy 241 EAAKSTKAASSTKFPDGFGLGEOIOWOAGTTPMIFPISLYLAGEVYNOSFRITIL 300
 Db 303 EAAKSTKAASSTKFPDGFGLGEOIOWOAGTTPMIFPISLYLAGEVYNOSFRITIL 362
 Qy 301 PQYLRPVEDVATSDQDCYKFAISQSSGTGVMGAVIMEGFVYVDFRARRKIRGAVASCHV 360
 Db 363 PQYLRPVEDVATSDQDCYKFAISQSSGTGVMGAVIMEGFVYVDFRARRKIRGAVASCHV 422
 Qy 361 H 361
 Db 423 H 423
 RESULT 2
 BACE-MOUSE STANDARD; PRT: 501 AA.
 ID AC P56818;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
 DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
 DE protease 2) (Asp 2) (Asp2) (Membrane-associated aspartic protease 2)
 DE (Memapsin-2).
 DE BACE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20002972; PubMed-10531052;
 RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
 RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
 RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
 RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
 RA Treanor J., Rogers G., Citron M.;
 RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.";
 RL Science 286:735-741(1999).
 RN [2]
 RP REVISIONS TO 6 AND 81-87.
 RA Bennett B.D., Vassar R., Citron M.;
 RA submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20057170; PubMed-10591213;
 RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
 RA Brashears J.R., Strahan N.C., Matthews W.R., Buhl A.E., Carter D.B.,
 RA Tomasselli A.G., Parodi L.A., Heinrichson R.L., Guney M.E.;
 RT "Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity.";
 RL Nature 402:533-537(1999).
 CC -1- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP, LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL


```

CC FRAGMENT WHICH IS LATER RELEASE BY GAMMA-SECRETASE (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: BRAIN
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF190726; AAF0143.2; -
CC EMBL: AF200346; AAF17082.1; -
CC HSSP: P56272; IAMS.
CC MGD: MGI:1346542; Bace.
CC InterPro: IPR001969; Asp-protease.
CC Pfam: PF000026; asp. 1.
CC PRINTS: PR00792; Pepsin.
CC PROSITE: PS00141; Asp_PROTEASE; 1.
CC KMW: Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
CC Signal.
CC FT SIGNAL 1 21 POTENTIAL.
CC FT PROPEP 22 45 POTENTIAL.
CC FT CHAIN 46 501 BETA-SECRETASE.
CC FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 458 478 POTENTIAL.
CC FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
CC FT ACT_SITE 93 93 BY SIMILARITY.
CC FT ACT_SITE 289 289 BY SIMILARITY.
CC FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 501 AA; 55747 MW; C085A013145E474E CRC64;

Query Match 99.4%; Score 1896; DB 1; Length 501;
Best Local Similarity 98.6%; Pred. No. 5.1e-152;
Matches 356; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```

RESULT 3
BACE_RAT
ID BACE_RAT STANDARD: PRT: 501 AA.
AC P56819;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
DE (Beta-site amyloid precursor protein cleaving enzyme) (aspartyl
DE protease 2) (asp 2) (asp2) (Membrane-associated aspartic protease 2)
DE (Memapsin-2).
DE BACE.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20002972; PubMed=10531052;
RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
RA Denis P., Teplow D.B., Ross S., Amaratne P., Loeloff R., Luo Y.,
RA Fisher S., Fuller J., Edenson S., Lile J., Jaroslinski M.A.,
RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
RA Treanor J., Rogers G., Citron M.;
RA "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
RT the transmembrane aspartic protease BACE."
RL Science 286:735-741(1999).
CC -1- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
CC AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
CC THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,
CC LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
CC SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
CC FRAGMENT WHICH IS LATER RELEASE BY GAMMA-SECRETASE (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF190727; AAF0144.1; -
CC HSSP: P32329; IYPS.
CC InterPro: IPR001969; Asp-protease.
CC Pfam: PF000026; asp. 3.
CC PRINTS: PR00792; Pepsin.
CC PROSITE: PS00141; Asp_PROTEASE; 1.
CC KMW: Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
CC Signal.
CC FT SIGNAL 1 21 POTENTIAL.
CC FT PROPEP 22 45 POTENTIAL.
CC FT CHAIN 46 501 BETA-SECRETASE.
CC FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 458 478 POTENTIAL.
CC FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
CC FT ACT_SITE 93 93 BY SIMILARITY.
CC FT ACT_SITE 289 289 BY SIMILARITY.
CC FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 501 AA; 55806 MW; 24B445BC8B87DE3 CRC64;

Query Match 99.1%; Score 1890; DB 1; Length 501;
Best Local Similarity 98.3%; Pred. No. 1.6e-151;
Matches 355; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 MYDNLRGSGGGYVYEWGSPQTLNIIIVDTGSSNFAVGAHPFLHRYROLSTYR 60
DB 63 MYDNLRGSGGGYVYEWGSPQTLNIIIVDTGSSNFAVGAHPFLHRYROLSTYR 122
OY 61 DLKRGVYPTTGKMGEGELGTDLVSIPIHGPNTVAVANIAITSDKFEFNGSNGEIGL 120
DB 123 DLKRGVYPTTGKMGEGELGTDLVSIPIHGPNTVAVANIAITSDKFEFNGSNGEIGL 182
OY 121 AVAELARPDDSLPEPPDLSLVKOTHPNLFSLDLCGAGPFLNGLSEVLASVSGSMITGIDH 180
DB 183 AVAELARPDDSLPEPPDLSLVKOTHPNLFSLDLCGAGPFLNGLSEVLASVSGSMITGIDH 242
OY 181 SLYTGSLWTPTRRREYVYVIVRVEINQDGMKDCKEYNYKSYDSCGTMLRLPKYF 240
DB 243 SLYTGSLWTPTRRREYVYVIVRVEINQDGMKDCKEYNYKSYDSCGTMLRLPKYF 302
OY 241 EAAVSIKASSTKPEPPDGLVLCQAGTTPWNIFPVISLYLMGEVNTQSPRITLL 300
DB 303 EAAVSIKASSTKPEPPDGLVLCQAGTTPWNIFPVISLYLMGEVNTQSPRITLL 362
OY 301 PQGYLRPVDVATSDDDCKFAISOSTGTVMGAVIMEGYVPEPRARRIGFAVSACHV 360
DB 363 PQGYLRPVDVATSDDDCKFAISOSTGTVMGAVIMEGYVPEPRARRIGFAVSACHV 422
OY 361 H 361
DB 423 H 423

RESULT 4
BAE2_HUMAN
ID BAE2_HUMAN STANDARD: PRT; 518 AA.
AC OGV520: OGVJ76:
DT 30-MAY-2000 (rel. 39, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Beta secretase 2 precursor (EC 3.4.23.-) (Beta-site APP-cleaving
DE enzyme 2) (Aspartyl protease 1) (Asp 1) (Asp1) (Membrane-associated
DE aspartic protease 1) (Memapsin-1).
GN BACE2 OR ASP21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057170; PubMed=10591213;
RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA Tomasselli A.G., Parodi L.A., Heinrichson R.L., Gurney M.E.;
RT "Membrane-anchored aspartyl protease with Alzheimer's disease
RT beta-secretase activity.";
RL Nature 402:533-537(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Xin H., Stephens J.C., Duan X., Harrowe G., Kim E., Grieshammer U.,
RA Giese K.;
RT "Identification of a novel aspartic-like protease differentially
RT expressed in human breast cancer cell lines.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Accorini M.P., Fumagalli P., Ottolenghi S., Taramelli R.;
RT "Cloning of a gene from chromosome 21 Down region encoding a potential
RT transmembrane aspartyl protease.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Solans A., Estivill X., de la Luna S.;
RT "Cloning of a novel mammalian aspartyl protease.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [5]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=20030166; PubMed=10561122;
RA Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
RA Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;
RT "Identification of a novel aspartic proteinase (Asp 2) as
RT beta-secretase.";
RL Mol. Cell. Neurosci. 14:419-427(1999).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=20144060; PubMed=10677483;
RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
RT "Human aspartic protease mepsin 2 cleaves the beta-secretase site of
RT beta-amyloid precursor protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Oht M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
RA Sharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehnemeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehnach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF200342; AAF17078.1; -
DR EMBL: AF117892; AAD45240.1; -
DR EMBL: AF050171; AAD45963.1; -
DR EMBL: AF178532; AAF28499.1; -
DR EMBL: AF204944; AAF26368.1; -
DR EMBL: AF200192; AAF13714.1; -
DR EMBL: AL163284; CAB90458.1; -
DR EMBL: AL163285; CAB90554.1; -
DR HSP: P00797; 2REN.
DR MIM: 605668; -
DR InterPro: IPR001969; Asp-protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF00026; asp; 3.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
KW Signal.
FT SIGNAL 1 20
FT PROPEP 21 ?
FT CHAIN ? 518
FT DOMAIN 21 473
FT TRANSMEM 474 494
FT DOMAIN 495 518
FT ACT_SITE 110 110
FT ACT_SITE 303 303
FT CARBOHYD 170 170
FT CARBOHYD 366 366
FT CONFLICT 36 A -> T (IN REF. 6).
FT SEQUENCE 518 AA; 56180 MW; 2E903150823760D3 CRC64;
SQ

```


PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE: 2.
 KW Hydrolyase; Aspartyl protease; Glycoprotein; Lysosome; Zymogen.
 FT PROPEP 1 44 ACTIVATION PEPTIDE.
 FT CHAIN 45 390 CATHEPSIN D.
 FT ACT_SITE 77 77
 FT ACT_SITE 77 77
 FT ACT_SITE 273 273
 FT ACT_SITE 71 140
 FT DISULFID 90 97
 FT DISULFID 264 268
 FT DISULFID 307 344
 FT DISULFID 114 114
 FT CARBOHYD 241 241
 FT SEQUENCE 390 AA; 4248 MW; 5B38A1C33C48D35 CRC64;
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Query Match 16.1%; Score 306.5; DB 1; Length 390;
 Best Local Similarity 28.1%; Pred. No. 2.1e-18;
 Matches 107; Conservative 69; Mismatches 118; Indels 87; Gaps 19;
 QY 13 YVEMTVGSPQTLNIIIVDGGSNFVAGAPHPFL-----HRYGROLSTYRDLRKGV 66
 DB 59 YVGEIGTPOCTVVFVFDIGSANDWPSHCKLIDLACWTHRKYNDSKSTY--VANGT 116
 QY 67 Y--VPYTGKMEGELGDLVSIPIHGPN-----VTVRANI--AAITESDKFFINGSMEG 116
 DB 117 TFDIHGSGSLGVLGQDTGVPNCNPSSSPGVTVQRTGGAIKQGVVFI--AAKEDG 175
 QY 117 ILGLAVAIAPRDSLEFPFSLVKQTHV--PNLFSLOLCAGFPINLOSEVLASVGGSMII 175
 DB 176 ILGMAYVPRIS--VNVNVLVEFDNLMOQKLVDRKVS-----FELNR--DPAQPGGELML 225
 QY 176 GGDHSLTGSIMWTPPIREWEYEVIIIVVEINGODLKMCKEYNDKSIYDSGTTNRL 235
 DB 226 GGTDSKYIRGSLFNFVNTROAWQIHMQDLV--GSSLIV--CK--GGCAIYDTGSLIVG 281
 QY 236 PKRYEAAVKSIRKAASSTKPEPDGFWLGEOLV--CWQAGTTPNNIPVLSILMGVTVMS 294
 DB 282 PVEEVRLEOKAIGAVPLIQ-----GEYMIPEKRVSS--LPEYTVKLG----- 322
 QY 295 FRITTLPOQLRPVEDVATNSODCYKFAISOSSTGT-----VMGAVI 336
 DB 323 -----KDYALSPED--YALKVSOAETTVCLSGFGMDIPPGGJMLIGDYF 367
 QY 337 MEGFYVEDRARRKIGFAVSA 357
 DB 368 IGRYTVFDRQNRVGLAEAA 388
 RESULT 7
 PEP1_RABIT STANDARD; PRT; 387 AA.
 AC P28712:
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Pepsin II-1 precursor (BC 3.4.23.1) (Pepsin A).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91009127; PubMed-2129536;
 RA Kagiyama T., Tanabe K., Koiwai O.;
 RT "Structure and development of rabbit pepsinogens. Stage-specific
 RT zymogens, nucleotide sequences of cDNAs, molecular evolution, and
 RT gene expression during development."
 RL J. Biol. Chem. 265:17031-17038(1990).
 CC -1- FUNCTION: SHOWS PARTICULAR BROAD SPECIFICITY, ALTHOUGH BONDS
 CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
 CC ALSO CLEAVED TO SOME EXTENT.
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE PHE-, LEU-
 CC -1- DEVELOPMENTAL STAGE: PEPSINOGENS IN GROUP I, II, AND III WHERE

THE PREDOMINANT ZYMGENS AT LATE POSTNATAL STAGE.
 CC -1- MISCELLANEOUS: THE EXPRESSION OF PEPsin GENES IS REGULATED BY
 CC HORMONES AND RELATED SUBSTANCES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 DR PTR, B38302; B38302.
 DR HSP; P00791; 1PSA.
 DR MEROPS; A01.001; -;
 DR InterPro; IPR001969; Asp-protease.
 DR InterPro; IPR001461; Pepsin.
 DR Pfam; PF00026; asp. 1.
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE: 2.
 KW Hydrolyase; Aspartyl protease; Digestion; Stomach; Gastric juice;
 KW Zymogen; Signal; Phosphorylation; Multigene family.
 FT SIGNAL 1 15
 FT PROPEP 16 59 ACTIVATION PEPTIDE.
 FT CHAIN 60 387 PEPsin II-1.
 FT MOD_RES 129 129 PHOSPHORYLATION (POTENTIAL).
 FT ACT_SITE 93 93 BY SIMILARITY.
 FT ACT_SITE 276 276 BY SIMILARITY.
 FT DISULFID 106 111 BY SIMILARITY.
 FT DISULFID 267 271 BY SIMILARITY.
 FT DISULFID 310 343 BY SIMILARITY.
 FT SEQUENCE 387 AA; 42070 MW; A6EC48F715541A48 CRC64;
 Query Match 16.0%; Score 306; DB 1; Length 387;
 Best Local Similarity 27.1%; Pred. No. 2.3e-18;
 Matches 98; Conservative 67; Mismatches 131; Indels 66; Gaps 15;
 QY 13 YVEMTVGSPQTLNIIIVDGGSNFVAG--AAHPFLHRYGROLSTYRDLRKGVV 68
 DB 75 YVGEIGTPOCTVVFVFDIGSANDWPSYTCSSLACFLHFRPDDSTFGQATSELSI 134
 QY 69 PYTGKMEGELGDLVSIPIHGPNTVRANIATIESD---KEFINGSNWESICGLAVAI 125
 DB 135 TYGGSMTGILGYFVKKV---GNIEDTQIPLGSKTEPGITFLV--APFDGILAVPSI 189
 QY 126 ARPDLSLEFPFDSLVKQTHV--PNLFSLOLCAGFPINLOSEVLASVGGSMIIGIDHSLYT 184
 DB 190 SASDAT--PVDNMMWNEGLVSEDLFVYLSNG-----EKGSVMVGGIDSSYIT 237
 QY 195 GSIWTPPIREWEYEVIIIVVEINGODLKM--DCKEYNDKSIYDSGTTNRLPKYFEA 242
 DB 238 GSLNVPVSHCYQIOTIMDSITINGETIACADSC-----QAVYDTGSLLAGFSAISK 291
 QY 243 AVKSIRKAASSTKPEPDGFWLGEOLV--CWQAGTTPNNIPVLSILMGVTVMSRITLP 301
 DB 292 IQSYIGASKNL-----LGENITCSAIDSLPDIVF-----TINN 325
 QY 302 QOYLRPVED--VATSDDC---YKFAISOSSTGT--VMGAVIMEGYVEDRARRKIGFAV 355
 DB 326 VQYPPASAYIILKEDDCLSGFDGKNLDTSTGELMILDGVLIRQYTFYFDRANNVGLAA 385
 QY 356 SA 357
 DB 386 AA 387
 RESULT 8
 PEP1_CHICK STANDARD; PRT; 367 AA.
 AC P00793:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Pepsin A precursor (EC 3.4.23.1).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]

```

RP SEQUENCE.
RX MEDLINE-84004412; PubMed-6617663;
RA Baudys M., Kostka V.;
RT "Covalent structure of chicken pepsinogen.";
RL Eur. J. Biochem. 136:89-99(1983).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS
CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC ALSO CLEAVED TO SOME EXTENT.
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE PHE-, LEU-.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
DR PIR: A00984; PEPH.
DR HSSP: P00794; 4CMS.
DR MEROPS: A01.001; -.
DR InterPro: IPR001969; Asp-protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF00026; asp. 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP-PROTEASE; 2.
KM Hydrolyase: Aspartyl protease; Digestion; Stomach; Zymogen;
KW Glycoprotein; Gastric juice.
FT PROPEP 1 42
FT CHAIN 43 367
FT ACT_SITE 77 77
FT ACT_SITE 260 260
FT CARBOHYD 113 113
FT DISULFID 90 95
FT DISULFID 251 255
FT DISULFID 290 323
SQ SEQUENCE 367 AA; 40431 MW; 0C547E7FD8F5B341 CRC64;

Query Match 15.8%; Score 302; DB 1; Length 367;
Best Local Similarity 24.0%; Pred. No. 4.7e-18;
Matches 88; Conservative 69; Mismatches 125; Indels 84; Gaps 13;

QY 13 YVEMTVGSPPTQTLNLTVDGSSNFVAGAPHPFL-----HRYQRQLSSTYRDLRG 65
DB 59 YGCTISIGTPQDFSVIFPDGSSNLWV---PSLYCKSSACSNHKKRPDPKSSSTYVSTNET 115
QY 66 VVYPYTGKMEGELGTDLVSIPIGPNTVYRANIAITESDK-FFINGSNMEGILGLAYAE 124
DB 116 VYIAYGTGSMGILGYDTAVV---SSIDVQNGIFGLSTEPGSEFFYCNFDGILGLAFPS 172
QY 125 IARDDSLPEFPDSLYKQTHV-PNLSLQLCGAGFPNLQSEVLASVGGSMITGIDHSILY 183
DB 173 IS--SSGATPPVEDNMMSQHLVADLESVYLSKDG-----ETGSEVLEGGIDPNYNT 220
QY 184 TQSLWTPPIRREYVEIVRVEINGODLK--MDCKEYNYDKSIYDSGTNLRKKVFE 241
DB 221 TKGIVWPLSAETWQITMDRVTGKRYVACFFTC-----QAIYDTGSLVLMFOGAIN 274
QY 242 AAVKSKAKASSNE-----KPPDGFWLGEOQLVCMQAGTTPWNIFFPVISLYLMEVTVQS 294
DB 275 RIITKDLGVSSDEISODDISKLDP-----VTFHNGHA----- 307
QY 295 FRITTLPOQYLRPEVDVATODCYKFAISQSSGT-----VMGAVIMEGGYVFPDAR 348
DB 308 -----FTLPARAVYINEGSCMLGFENMGTPTELGEOWILGDAVETREYVIFDRAN 358
QY 349 KRIGFA 354
DB 359 NKVGLS 364

```

```

OS Macaca fuscata fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
ON NCB1_TaxID=9543;
RX TISSUE-gastric mucosa;
RX MEDLINE-92037645; PubMed-1935977;
RA Kagayama T., Tanabe K., Koizumi O.;
RT "Development-dependent expression of isozymogens of monkey
RT pepsinogens and structural differences between them.";
RL Eur. J. Biochem. 202:205-215(1991).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS
CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC ALSO CLEAVED TO SOME EXTENT.
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE PHE-, LEU-.
CC -1- MISCELLANEOUS: THE EXPRESSION OF PEPsinogen GENES IS REGULATED BY
CC HORMONES AND RELATED SUBSTANCES.
CC -1- MISCELLANEOUS: EACH PEPsinogen IS CONVERTED TO CORRESPONDING
CC PEPsin AT PH 2.0 IN PART AS A RESULT OF THE RELEASE OF A 47 AA
CC CLEAVAGE VIA AN INTERMEDIATE FORM(S).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59753; CAA42425.1; -.
DR PIR: S16065; S16065.
DR PIR: S19682; S19682.
DR HSSP: P00790; IPSN.
DR MEROPS: A01.001; -.
DR InterPro: IPR001969; Asp-Protease.
DR InterPro: IPR001791; Laminin-G.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF00026; asp. 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP-PROTEASE; 2.
KW Hydrolyase; Aspartyl protease; Digestion; Stomach; Gastric juice;
KW Zymogen; Multigene family; Signal; Glycoprotein.
FT SIGNAL 1 15
FT PROPEP 16 38
FT PROPEP 39 62
FT CHAIN 63 388
FT ACT_SITE 94 94
FT ACT_SITE 277 277
FT DISULFID 107 112
FT DISULFID 268 272
FT DISULFID 311 344
FT CARBOHYD 88 88
SQ SEQUENCE 388 AA; 41955 MW; A2923AB1F7CDEB9 CRC64;

Query Match 15.8%; Score 302; DB 1; Length 388;
Best Local Similarity 28.6%; Pred. No. 5.1e-18;
Matches 103; Conservative 58; Mismatches 131; Indels 68; Gaps 15;

QY 13 YVEMTVGSPPTQTLNLTVDGSSNFVAGAPHPFL-----HRYQRQLSSTYRDLRG 65
DB 76 YFGTIGITPAQNFTVYFDGSSNLWV---PSYCYSLACMDNHLNPPDSSSTYRATSKT 132
QY 66 VVYPYTGKMEGELGTDLVSIPIGPNTVYRANIAITESDK-FFINGSNMEGILGLAYAE 124
DB 133 VSIYTGSMGILGYDTAVV---GGISDINQIFGLSTEPGFLFAFPDGLGLAFPS 189
QY 125 IARDDSLPEFPDSLYKQTHV-PNLSLQLCGAGFPNLQSEVLASVGGSMITGIDHSILY 183

```


CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84424; AAA52300.1; -
DR EMBL; M84413; AAA52300.1; JOINED.
DR EMBL; M84417; AAA52300.1; JOINED.
DR EMBL; M84418; AAA52300.1; JOINED.
DR EMBL; M84419; AAA52300.1; JOINED.
DR EMBL; M84420; AAA52300.1; JOINED.
DR EMBL; M84421; AAA52300.1; JOINED.
DR EMBL; M84422; AAA52300.1; JOINED.
DR EMBL; J05036; AAA52300.1; -
DR EMBL; AJ250717; CAB82850.1; -
DR PIR; A34401; A34401.
DR PIR; A34643; A34643.
DR PIR; A42038; A42038.
DR HSSP; P00794; 4CMS.
DR MEROPS; A01.010; -
DR MIM; I16890; -
DR InterPro; IPR001969; Asp-protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp.1.
DR PRINTS; PR00792; Pepsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolyase; Aspartyl protease; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 17
FT PROPEP 18 53
FT CHAIN 54 396
FT MOD_RES 18 18
FT ACT_SITE 96 96
FT ACT_SITE 281 281
FT DISULFID 60 60
FT DISULFID 109 114
FT DISULFID 272 276
FT DISULFID 314 351
FT CARBOHYD 90 90
FT CARBOHYD 220 220
FT CARBOHYD 333 333
SQ SEQUENCE 396 AA; 42793 MW; 40B643C5FB01521E CRC64;

Query Match 15.7%; Score 298.5; DB 1; Length 396;
Best Local Similarity 26.7%; Pred. No. 1e-17;
Matches 96; Conservative 61; Mismatches 141; Indels 61; Gaps 14;

OY 13 YVEKTVGSPPTLNLIVDTGSSNFAYGA---APHPLHRYRQLSTYRDLRKGVYV 68
DB 78 YFGRTISGTPQDFYVIFDTGSSNLWVSVCTSPACKTHSRFPQSGSTYSQGGQSSSI 137
OY 69 PYTGKMEGELGTDVLSVPHRPNTVTRANIAITESDKFFINGSNWEGITGLAYAEIARP 128
DB 138 OYGTGSLSGITIGADGVSV-BELIVYGGQFGSGSVTEPGTFVD-AEFDGILGLGYPSLA-- 193
OY 129 DQSLPEPFDLSVKTQHVNLPSLQCGAGFPLNQSEVLAVSGVSMIGGIDHSITYSGLW 188
DB 194 VGGVTPVPEDDNMAQ---NLVDLPMFSVYMSNPE---GAGSGELIIGGIDHSIFSSSLN 246
OY 189 YTPIRREMYEVIIIVRVLNQLDKMCKEYNYKSYVDSGTNLRLPKVFEAAVKSIR 248
DB 247 WVPYTKAWIWDIADNLQVGG--TVMFCS--GCAIYDGTSLITGSPDKIKQLQNAIG 302
OY 249 AASSTKEPDEFMLGELQVLCWAGCTPMNIPVSLVLMGEVTVNQSRITLLPQOYLAPV 308
DB 303 AAP-----VDESEVAVE-----CANLWAPDVTFTING-----VPYLSPTAY--TL 341

OY 309 EDVATNSQDDCYKFAISQSTG-----TWGAVIMEGYVVEPRARRRIGFA 354
DB 342 LDFVDGMFC-----SSGFQGLDHPHPAGPLMILGDVFIROPYSVDFRGNRVGLA 392
RESULT 12
ID PEPT_RABIT STANDARD: PRT: 387 AA.
AC P27821.
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NM NCI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91009127; PubMed=2129536;
RA Kageyama T., Tanabe K., Koizumi O.;
RT "Structure and development of rabbit pepsinogens. Stage-specific
RT zymogens, nucleotide sequences of cDNAs, molecular evolution, and
RT gene expression during development.";
RL J. Biol. Chem. 265:17031-17038(1990).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS
CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC ALSO CLEAVED TO SOME EXTENT.
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE PHE-, LEU-
CC -1- DEVELOPMENTAL STAGE: PEPSINOGENS IN GROUP I, II, AND III WHERE
CC THE PREDOMINANT ZYMOGENS AT LATE POSTNATAL STAGE.
CC -1- MISCELLANEOUS: THE EXPRESSION OF PEPSINOGEN GENES IS REGULATED BY
CC HORMONES AND RELATED SUBSTANCES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59235; AAA83369.1; -
DR PIR; C38302; C38302.
DR HSSP; P00790; 1PSN.
DR MEROPS; A01.001; -
DR InterPro; IPR001969; Asp-protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp.1.
DR PRINTS; PR00792; Pepsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolyase; Aspartyl protease; Digestion; Stomach; Gastric juice;
KW Zymogen; Signal; Phosphorylation; Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 59
FT CHAIN 60 387
FT MOD_RES 129 129
FT ACT_SITE 93 93
FT ACT_SITE 276 276
FT DISULFID 106 111
FT DISULFID 267 271
FT DISULFID 310 343
SQ SEQUENCE 387 AA; 42100 MW; 66FC331A3DC75891 CRC64;

Query Match 15.5%; Score 296; DB 1; Length 387;
Best Local Similarity 26.9%; Pred. No. 1.6e-17;
Matches 97; Conservative 63; Mismatches 135; Indels 66; Gaps 13;

OY 13 YVEKTVGSPPTLNLIVDTGSSNFAYGAAPHF-----LHRYRQLSTYRDLRK 65
DB 75 YFGRTISGTPQDFYVIFDTGSSNLWV---PSTYCSIALCALHKKRNPEDSSTYGTSET 131

```

QY 66 YVVPYTGKWEGLCTDVLSPHGPNTVRANIAITESDKFETNGSWEGLGLAYAEI 125
DB 132 LSTYGTGSMGILCYDVKVKSIEDTQIRGLSTEPSLFLF--APFDGILLGAPSI 189
QY 126 ARPDLSLPEFDSLVKQTHV-PNLFSLQCGAGPPLNOSVLAASVGSMITGGIDHSIXT 184
DB 190 SSSDAT--PVFDNMNMNEGVLSDLFVYLSSD-----EKGSLVMFGGIDSSYTT 237
QY 185 GSWMTPIRREKYYEVILVREINGODLKM--DCKEYNDKSYDSTGNTLRPKVVEA 242
DB 238 GSLNMVPVSYEYQWITMDVSINGETIACADSC-----QAIYDTGSLTGP-----TS 287
QY 243 AVKSITKAASSTKFPDGLGEOLV-CWQAGTTPWNIFPVISLYLMGEVYNOSFRITLLP 301
DB 288 AISNIOSTYGAK-----NLGENVISCSAIDSLPDIIV-----TING 325
QY 302 QOYLPRVEDVATSDDCYKFAISQSTG-----YMGAVIMEGYVEDRARKRIGFAV 355
DB 326 IQYPLPASAVILKEDDCTSGLEGNMVDITYGELMIGDVFIRQYFYFDRANNOGLAA 385
QY 356 S 356
DB 386 A 386

RESULT 13
CATD_CAVPO STANDARD: PRT: 391 AA.
AC P25796;
DT 01-MAY-1992 (rel. 22, Created)
DT 01-MAY-1992 (rel. 22, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Cathepsin E precursor (EC 3.4.23.34).
GN CTSE.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniala; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OC NCB1_Taxid=10141;
RN (1)
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=92355614; PubMed=1644829;
RA Kageyama T., Ichinose M., Tsukada S., Miki K., Kurokawa K., Koiwai O.,
RA Tanji M., Yakabe E., Athuda S.B., Takahashi K.;
RT "Gastric procathepsin E and progastricins from guinea pig.
RT Purification, molecular cloning of cDNAs, and characterization of
RT enzymatic properties, with special reference to procathepsin E.";
RL J. Biol. Chem. 267:16450-16459(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE-Gastric mucosa;
RX MEDLINE=96073637; PubMed=8540321;
RA Kageyama T., Ichinose M., Miki K., Moriyama A., Yonezawa S.,
RA Tanji M., Athuda S.B., Takahashi K.;
RT "Isolation, characterization, and structure of procathepsin E and
RT cathepsin E from the gastric mucosa of guinea pig.";
RL Adv. Exp. Med. Biol. 362:211-221(1995).
CC -1- FUNCTION: DUE TO ITS INTRACELLULAR LOCATION AND DISTRIBUTION IN
CC LYMPHOID ASSOCIATED TISSUE, IT MAY HAVE A ROLE IN IMMUNE FUNCTION.
CC -1- CATALYTIC ACTIVITY: Similar to cathepsin D, but slightly broader
CC specificity.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: M88653; AAA37052.1; -.
DR EMBL: S80547; AAB35844.1; -.
DR PIR: A43356; A43356.
DR HSSP: P00794; 4CMS.
DR MEROPS: A01.010; -.
DR InterPro: IPR001969; Asp-protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF000026; asp.1.
DR PRINTS: PRO0792; PEPSTN.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
KW Hydrolyase; Aspartyl protease; Glycoprotein; zymogen; signal.
FT SIGNAL 1 18
FT PROPEP 19 53
FT CHAIN 54 391
FT ACT_SITE 92 92
FT ACT_SITE 276 276
FT DISULFID 56 56
FT DISULFID 105 110
FT DISULFID 267 271
FT DISULFID 309 346
FT CARBOHYD 86 86
SQ SEQUENCE 391 AA; 42132 MW; 78D216BF8C6DABD CRC64;

Query Match 15.5%; Score 296; DB 1; Length 391;
Best Local Similarity 26.9%; Pred. No. 1,6e-17;
Matches 98; Conservative 64; Mismatches 130; Indels 72; Gaps 16;

QY 13 YVENVGSSPQTLNIIVDTGSSNPAVGA---APHPFLHRYRORLSSTYRDLKRGVYV 68
DB 74 YFGTISGSPQNTFYVDTGSSNLMWPSVYCTSPACQTHVFPRLSSTYRRENGSFSI 133
QY 69 PYTGKWEGLCTDVLSPHGPNTVRANIAITESDKFETNGSWEGLGLAYAEIARP 128
DB 134 QYGTGSLGIIIGADQVSV-EGLTIVGQGFGEVQEPKTFVH-AEFDGILLGAPSLAA- 190
QY 129 DLSLPEFDSLVKQTHVNPISLQCGAGPPLNOSVLAASVGSMITGGIDHSIYTSIM 188
DB 191 -GGVTPVFDNMMAQ---NLVALPM---FSYMSNPNNGSGSELTFGCIPIPSHFGSLN 241
QY 189 YPIRREMYEYIIVREINGODLMDCKEYNDKSYDSTGNTLRPKVFEAAVKSIR 248
DB 242 WVPVYKQAWQALDGIQVQ--DSVMFCSE--CCQAIYDTGSLITGP-----PKIKOLQ 293
QY 249 AASSTKFPDGLGEOLVQWQAGTTPWNIFPVISLYLMGEVYNOSFRITLLPQ 303
DB 294 EALGATVYDEGV---SVQC-----ANLMMMLDVT---FLINGPYTLNPTA 333
QY 304 YLRPVEDVATSDDCYKFAISQSTG-----TWGAVIMEGYVEDRARKR 350
DB 334 Y--TLIDFVDGMQVC-----STGFELQLTPPAGPLMIGDVFIRQYFVDFRGNNR 383
QY 351 IGFA 354
DB 384 VGIA 387

RESULT 14
CATD_HUMAN STANDARD: PRT: 412 AA.
ID CATD_HUMAN P07339;
AC P07339;
DT 01-APR-1988 (rel. 07, Created)
DT 01-APR-1988 (rel. 07, Last sequence update)
DT 01-MAR-2002 (rel. 41, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
GN CTSD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OC NCB1_Taxid=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=85270436; PubMed=3927292;
RA Faust P.L., Kornfeld S., Chirgwin J.M.;

```


RT "Cloning and sequence analysis of cDNA for human cathepsin D.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4910-4914(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87231068; PubMed=3588310;
 RA Westley B.R., May F.E.B.;
 RT "Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive
 human breast cancer cells.";
 RL Nucleic Acids Res. 15:3773-3786(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91299158; PubMed=2069717;
 RA Redeker B., Heckendorf B., Grosch H.W., Wersmann G., Hasilik A.;
 RT "Molecular organization of the human cathepsin D gene.";
 RL DNA Cell Biol. 10:423-431(1991).
 RN [4]
 RP SEQUENCE OF 1-22 FROM N.A.
 RX MEDLINE=94085791; PubMed=8262386;
 RA May F.E., Smith D.J., Westley B.R.;
 RT "The human cathepsin D-encoding gene is transcribed from an estrogen-
 regulated and a constitutive start point.";
 RL Gene 134:277-282(1993).
 RN [5]
 RP SEQUENCE OF 1-22 FROM N.A.
 RX MEDLINE=95021301; PubMed=7935485;
 RA Augereau P., Miralles F., Cavailles V., Gaudeliet C., Parker M.,
 RT Rochefort H.;
 RT "Characterization of the proximal estrogen-responsive element of
 human cathepsin D gene.";
 RL Mol. Endocrinol. 8:693-703(1994).
 RN [6]
 RP SEQUENCE OF 170-180.
 RC TISSUE=Liver;
 RA Hochstrasser D.F., Fritiger S., Paquet N., Bairoch A., Ravier F.,
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
 RA Appel R.D., Hughes G.J.;
 RT Submitted (JUN-1992) to the SWISS-PROT data bank.
 RN [7]
 RP VARIANT VAL-58.
 RX MEDLINE=20179010; PubMed=10716266;
 RA Pappasotiropoulos A., Bagli M., Kurz A., Kornhuber J., Forstl H.,
 RA Maier W., Pauls J., Lautenschlager N., Heun R.;
 RT "A genetic variation of cathepsin D is a major risk factor for
 Alzheimer's disease.";
 RL Ann. Neurol. 47:399-403(2000).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RC TISSUE=Spleen;
 RX MEDLINE=9323670; PubMed=8467789;
 RA Metcalf P., Fusek M.;
 RT "Two crystal structures for cathepsin D: the lysosomal targeting
 signal and active site.";
 RL EMBO J. 12:1293-1302(1993).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RC TISSUE=Liver;
 RX MEDLINE=93342076; PubMed=8393577;
 RA Baldwin E.T., Bhat T.N., Gulnik S., Hosur M.V., Sowder R.C. II,
 RA Cachau R.E., Collins J., Silva A.M., Erickson J.W.;
 RT "Crystal structures of native and inhibited forms of human cathepsin
 D: Implications for lysosomal targeting and drug design.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6796-6800(1993).
 CC -1- FUNCTION: CATHEPSIN D IS AN ACID PROTEASE ACTIVE IN INTRACELLULAR
 CC PROTEIN BREAKDOWN.
 CC -1- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
 CC that of pepsin A. Does not cleave the 4-Gln-I-His-5 bond in B
 CC chain of insulin.
 CC -1- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
 CC -1- SUBCELLULAR LOCATION: Lysosomal.
 CC -1- POLYMORPHISM: The Val-58 allele is significantly overrepresented
 CC in demented patients (11.8%) compared with nondemented controls
 CC (4.9%). Carriers of the Val-58 allele have a 3.1-fold increased
 CC risk for developing AD than noncarriers.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1. ALSO KNOWN AS THE
 CC EURARYOTIC ASPARTYL PROTEASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M11233; AAB59529.1; -
 DR EMBL; X05344; CA828855.1; -
 DR EMBL; M63138; AAB51922.1; -
 DR EMBL; M63134; AAB51922.1; JOINED.
 DR EMBL; M63135; AAB51922.1; JOINED.
 DR EMBL; M63136; AAB51922.1; JOINED.
 DR EMBL; M63137; AAB51922.1; JOINED.
 DR EMBL; L12980; AAB16314.1; -
 DR EMBL; S74689; AAD14456.1; -
 DR EMBL; S52557; AAD13668.1; -
 DR PIR; A25771; KHHUD.
 DR PDB; 1LYA; 31-JAN-94.
 DR PDB; 1LYB; 31-JAN-94.
 DR MEROPS; A01.009; -
 DR SWISS-2DPAGE; P07339; HUMAN.
 DR Siens-2DPAGE; P07339; -
 DR MIM; 116840; -
 DR InterPro; IPR001969; Asp-protease.
 DR InterPro; IPR001461; Pepsin.
 DR Pfam; PF00026; asp. 1.
 DR PRINTS; PRO0792; PEPsin.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Signal; zymogen;
 KW Polymorphism; Alzheimer's disease; 3D-structure.
 FT SIGNAL 1 18
 FT PROPEP 19 64 ACTIVATION PEPTIDE.
 FT CHAIN 65 412 CATHEPSIN D.
 FT CHAIN 65 161 LIGHT CHAIN (PROBABLE).
 FT ACT_SITE 169 412 HEAVY CHAIN (PROBABLE).
 FT ACT_SITE 97 97
 FT DISULFID 91 160
 FT DISULFID 110 117
 FT DISULFID 286 290
 FT DISULFID 329 366
 FT CARBOHYD 134 134
 FT CARBOHYD 263 263
 FT VARIANT 58 58
 FT STRAND 67 74
 FT TURN 75 77
 FT STRAND 78 85
 FT TURN 86 89
 FT STRAND 90 97
 FT TURN 98 99
 FT STRAND 103 107
 FT TURN 108 109
 FT TURN 112 113
 FT HELIX 115 118
 FT TURN 119 119
 FT STRAND 123 123
 FT HELIX 125 127
 FT STRAND 129 130
 FT TURN 132 141
 FT STRAND 146 158
 FT STRAND 172 184
 FT HELIX 188 192
 FT STRAND 197 200
 FT HELIX 204 206
 N-LINKED (GLCNAC. . .).
 A -> V (ASSOCIATED WITH INCREASED RISK IN
 AD; POSSIBLY INFLUENCES SECRETION AND
 INTRACELLULAR MATURATION; IN
 DBSNP:175711.
 /FTID=VAR_011621.

FT	HELIX	208	210
FT	HELIX	214	220
FT	TURN	221	222
FT	STRAND	228	233
FT	STRAND	243	247
FT	TURN	248	248
FT	HELIX	252	254
FT	STRAND	255	263
FT	STRAND	267	267
FT	TURN	268	269
FT	STRAND	270	279
FT	TURN	280	281
FT	STRAND	284	285
FT	TURN	287	288
FT	STRAND	290	294
FT	TURN	296	297
FT	STRAND	298	298
FT	STRAND	301	303
FT	HELIX	305	315
FT	TURN	316	316
FT	STRAND	318	319
FT	TURN	322	323
FT	STRAND	325	328
FT	HELIX	329	334
FT	STRAND	338	342
FT	TURN	343	344
FT	STRAND	345	349
FT	HELIX	351	354
FT	STRAND	355	357
FT	TURN	359	362
FT	STRAND	365	368
FT	STRAND	370	372
FT	TURN	377	379
FT	STRAND	383	385
FT	HELIX	387	390
FT	TURN	391	392
FT	STRAND	393	398
FT	TURN	399	402
FT	STRAND	403	409

Query Match	15.5%	Score 295.5	DB 1	Length 412
Best Local Similarity	28.5%	Pred. No. 1.9e-17		
Matches 105	Conservative 58	Mismatches 146	Indels 59	Gaps 15
Qy	13	YYVEMTVGSPQPLNLIVDTGSSNFAGAPHPFL-----HRYQROLSTYRDLKGV	66	
Db	79	YGEIGTGTPOCFYTFVDGSSNLWPSLHCKLLDIACIHHKYNDSKSTFYVKNQTS	138	
Qy	67	YVPYTGKWEGLGTDLVIP-----HGNVTVRANIAAITESDKFFINGSNWEG	117	
Db	139	DIHVGSSLSGLYSDQTVSPVCSASSASALGKVYKROVERGEATKOPGITFAARFDG	198	
Qy	118	LGLAATLARPDDSLPEPFIYVQKTHV-PNLESLQCGAGFPLNOSSEVLASVGSMTIG	176	
Db	199	LGMAYPRIS-VNNVLFVFDNLMOQKVLQDNIISFYL-----SNDPAQGGGELMTG	248	
Qy	177	GIDSLYTGSLMTPYRIRREWYEVIIIVREI-NGQDLKMDCKEYNDKSYDSGTMLRL	235	
Db	249	GTDCKYKYSLSYINVTYRKAAWQYHLLQVEVASGLTL--CKE--GCEALVDGTSLMWG	303	
Qy	236	PKKYPEAIVKSIKKAASSTENKPPDFWIMGEOLV-CMOAGTTPMNIFPYISLYLMGEVYNQS	294	
Db	304	PVDEVRLELQKAIQAVPLIQ-----GEYMIPEKYST-----LPALITLKG-----KG	346	
Qy	295	FRITILLPOQYLRPEVDVATSQDDCYKFAISO-----SSTGTVMGAVIMEGYVYFDPARK	349	
Db	347	YKLS--PEDYTLTKYSAQKTL--CISGPMGMDIPPPSGPLTWILGDVIFIGRAYTFVDRDNN	402	
Qy	350	RIGFAVSA 357		
Db	403	RVGFEAAA 410		

[illegible]

Db 282 LTGP---TSAISNISYIGASK---NLGENVISCSAIDSLPDIVF----- 321
QY 292 NOSFRITILPOQYLRPVEDVATSODDCYKFAISOSSTGT-----VMGAVIMEGFYVFD 345
Db 322 -----TINGIQYPLPASAVILKEDDDCTSGLEGMMVDYTGELMILGDVFIQYFTVFD 375
QY 346 RARRRIGFAYS 356
Db 376 RANNQGLLAAA 386

Search completed: October 30, 2002, 12:27:48
Job time : 32.0442 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:30 ; Search time 112.646 seconds
(without alignments)
554,401 Million cell updates/sec

Title: US-09-724-571-75

Perfect score: 1907
Sequence: 1 MVDNLRKSGGGYIVEMTVG.....VVEDRRKRKIGFAVSACHVH 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1907	100.0	532	4 Q9ULS1	Q9ULS1 homo sapien
2	1752.5	91.9	476	4 Q9BYC1	Q9BYC1 homo sapien
3	1649	86.5	457	4 Q9BYC0	Q9BYC0 homo sapien
4	1504.5	78.9	432	4 Q9BYB9	Q9BYB9 homo sapien
5	1106	58.0	439	4 Q9H2V8	Q9H2V8 homo sapien
6	1088	57.1	514	11 Q9JLI8	Q9JLI8 mus musculi
7	996	52.2	266	11 Q9CUD5	Q9CUD5 mus musculi
8	956.5	50.2	396	4 Q9NZL1	Q9NZL1 homo sapien
9	907	47.6	468	4 Q9NZL2	Q9NZL2 homo sapien
10	712.5	37.4	213	4 Q9PDD2	Q9PDD2 homo sapien
11	547.5	28.7	255	11 Q9RIP7	Q9RIP7 mus musculi
12	332.5	17.4	391	5 Q9VKP6	Q9VKP6 drosophila
13	332	17.4	354	5 Q9GIX7	Q9GIX7 boophilus m
14	308	16.2	384	13 Q9DEC2	Q9DEC2 xenopus lae
15	308	16.2	385	13 Q9DEC4	Q9DEC4 rana catesb
16	305	16.0	386	6 Q9GM77	Q9GM77 rhinolophus

17	305	16.0	387	6 Q9GM78	Q9GM78 scorax ungui
18	304.5	16.0	372	5 Q9VLK3	Q9VLK3 drosophila
19	304.5	16.0	386	6 Q9BGS5	Q9BGS5 bos taurus
20	304	15.9	387	6 Q9GM79	Q9GM79 suncus muri
21	304	15.9	387	13 Q9DDV5	Q9DDV5 salvelinus
22	303	15.9	383	13 Q9DEC3	Q9DEC3 xenopus lae
23	303	15.9	383	13 Q9DE45	Q9DE45 salvelinus
24	301	15.8	382	13 Q9PRG9	Q9PRG9 gallus gall
25	299	15.7	423	5 Q9VKP7	Q9VKP7 drosophila
26	298	15.6	376	13 Q9PDR8	Q9PDR8 pseudopleur
27	295.5	15.5	384	13 Q9I322	Q9I322 rana catesb
28	294	15.4	396	6 Q9J428	Q9J428 chionodraco
29	290.5	15.2	381	6 Q9GK11	Q9GK11 camelus dro
30	289	15.2	386	6 Q9GM76	Q9GM76 canis famli
31	287	15.0	399	13 Q9J458	Q9J458 canis famli
32	285.5	15.0	444	5 Q21966	Q21966 caenorhabdi
33	284	14.9	398	13 P87370	P87370 oncorhynch
34	280.5	14.7	427	5 P91802	P91802 schistosoma
35	279.5	14.7	390	6 Q9GK10	Q9GK10 camelus dro
36	278	14.6	370	6 Q9TWL1	Q9TWL1 bos taurus
37	278	14.6	378	13 Q9PDR9	Q9PDR9 pseudopleur
38	278	14.6	399	13 Q9DD89	Q9DD89 brachydanio
39	278	14.6	422	5 Q96906	Q96906 onchocerca
40	277.5	14.6	380	6 Q28950	Q28950 sus scrofa
41	276	14.5	396	13 Q9DEX3	Q9DEX3 clupea hare
42	275	14.4	385	6 Q29080	Q29080 sus scrofa
43	275	14.4	413	3 Q14413	Q14413 pichia angu
44	275	14.4	446	5 Q9N9H3	Q9N9H3 necator ame
45	272	14.3	389	13 Q9PWK1	Q9PWK1 gallus gall

ALIGNMENTS

RESULT 1

ID Q9ULS1 PRELIMINARY; PRT; 532 AA.
AC Q9ULS1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE KIAA1149 PROTEIN (FRAGMENT).
GN KIAA1149.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20039618; PubMed=10574461;
RA Hirosewa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;
RT "Characterization of cDNA clones selected by the Genemark analysis
from size-fractionated cDNA libraries from human brain.";
RL DNA Res. 6:329-336(1999).
DR EMBL; AB032975; BAA86463.2; -
DR HSSP; P56272; IAMS.
DR MEROPS; A01.004; -
DR InterPro; IPR001969; Asp-protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 532 AA; 58720 MW; 98B13500DSFBDZEB CRC64;

Query Match 100.0%; Score 1907; DB 4; Length 532;

Best Local Similarity 100.0%; Pred. No. 6e-159;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MVDNLRKSGGGYIVEMTVGSPQTLNIIVDTGSSNFAYGAAPHLRHRYQRQLSSYTR 60
Db 94 MVDNLRKSGGGYIVEMTVGSPQTLNIIVDTGSSNFAYGAAPHLRHRYQRQLSSYTR 153

OY	61	DLRGNYVPTYOGKMBEGLCTDLVSLPHGNNVVRANIAITAITSDDKFLNGSMWBEILGL	120
Db	154	DLRGVYVPTYOGKMBEGLCTDLVSLPHGNNVVRANIAITAITSDDKFLNGSMWBEILGL	213
OY	121	AYAIARPDOSLEFPFSLVKQTHVPLFSLQCGAGFPLNQSEVLASVGSMTIGIDH	180
Db	214	AYAIARPDOSLEFPFSLVKQTHVPLFSLQCGAGFPLNQSEVLASVGSMTIGIDH	273
OY	181	SLYGSLSMYPIRREMYEYVLIYVEINQDLMKDKKEYNDKSIYDSGTTNLRLEKYE	240
Db	274	SLYGSLSMYPIRREMYEYVLIYVVEINQDLMKDKKEYNDKSIYDSGTTNLRLEKYE	333
OY	241	EAAYKSIKAASSTKFPDGEFWLGSOLVCWQAGTTPNNIFPVISLYLMEVTNOSFRTIL	300
Db	334	EAAYKSIKAASSTKFPDGEFWLGSOLVCWQAGTTPNNIFPVISLYLMEVTNOSFRTIL	393
OY	301	POOYLRPEVEDYATSDDCYKRAISQSSGTVMGAVINMEGFYVDRRKRRIGFVAVSACHV	360
Db	394	POOYLRPEVEDYATSDDCYKRAISQSSGTVMGAVINMEGFYVDRRKRRIGFVAVSACHV	453
OY	361	H 361	
Db	454	H 454	

RESULT 2					
ID	O9BYC1	PRELIMINARY;	PRT:	476 AA.	
AC	O9BYC1:				
Dt	01-JUN-2001 (TrEMBLrel. 17, Created)				
Dt	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)				
Dt	01-DEC-2001 (TrEMBLrel. 19, last annotation update)				
DE	BETA-SITE APP CLEAVING ENZYME I-476.				
GN	BACE.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-BRAIN:				
RX	MEDLINE=21408467; PubMed=11516562;				
RA	Tanahashi H., Tabita T.;				
RT	"three novel alternatively spliced isoforms of the human beta-site APP				
RT	cleaving enzyme (BACE) and their effect on amyloid beta-peptide				
RT	production.";				
RL	Neurosci. Lett. 307:9-12(2001).				
DR	EMBL; AB050436; BAB40931.1; -.				
DR	HSSP; P32329; IYPS.				
DR	InterPro: IPRO01969; Asp.protease.				
DR	InterPro: IPR001461; Pepsin.				
DR	Pfam: PF00026; asp.1				
DR	PRINTS: PR00792; PEPSIN				
DR	PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1				
SQ	SEQUENCE 476 AA; 52907 MW; 6C8CB7E8A953AF66 CRC64;				
Query Match		91.9%; Score 1752.5; DB 4; Length 476;			
Best local similarity		93.1%; Pred. No.1.9e-145;			
Matches	336; Conservative %	0; Mismatches	0; Indels	25; Gaps	
Oy	1 MDNLNGKGGGCGYYVEMTVGSPQTILNYDTGSNFVAAGAPHLPHRYOROLSTSYR	60			
Db	63 MDNLNGKGGGCGYYVEMTVGSPQTILNYDTGSNFVAAGAPHLPHRYORQLSTSYR	122			
Oy	61 DLRKGYVVYITGKGWELGTDLIVSIPIGPNTVYRANIAAITESDKFFINGSNWGIIGL	120			
Db	123 DLKRGYVVYITGKGWELGTDLIVSIPIGPNTVYRANIAAITESDKFFINGSNWGIIGL	182			
Oy	121 AYAEIARPPDLSPEPFDSLAKOTHPNFLFSIQLGAGFPPLNQSVLASVGSGMIIGSDH	180			
Db	183 AYAEIAR-----LCGAGPPLNQSVLASVGSGMIIGSDH	217			

Qy	181	SLYGTSLWYTPIRREMYEVIIVAVEINQDLDKMDCKEYVYKDSIYDSGTTNLRPKKVE	240
Db	218	SLYGTSLWYTPIRREMYEVIIVAVEINQDLDKMDCKEYVYKDSIYDSGTTNLRPKKVE	277
Qy	241	EAAKSIKAASSTEKEFPDGFWLGSQLCWOAGTTPMNIIPVYISLYLMEVYVNSFRITIL	300
Db	278	EAAKSIKAASSTEKEFPDGFWLGSQLCWOAGTTPMNIIPVYISLYLMEVYVNSFRITIL	337
Qy	301	POQYLRPVEDVATSDDCYKFAISQSSTGYVMGAVIMEGFVYVDRKKRIIGFAVSCHV	360
Db	338	POQYLRPVEDVATSDDCYKFAISQSSTGYVMGAVIMEGFVYVDRKKRIIGFAVSCHV	397
Qy	361	H 361	
Db	398	H 398	

ID	Q9BYC0	PRELIMINARY:	PRT:	457 AA.
AC	Q9BYC0:			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	BETA-SITE APP CLEAVING ENZYME I-457 (BETA-SITE APP CLEAVING ENZYME TYPE C).			
GN	BACE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRNIN;			
RX	MEDLINE=21408467; PubMed=11516562;			
RA	Tanahashi H., Tabira T.;			
RT	"Three novel alternatively spliced isoforms of the human beta-site APP			
RT	cleaving enzyme (BACE) and their effect on amyloid beta-peptide			
RT	production.";			
RL	Neurosci. Lett. 307:9-12(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=EXOCRINE PANCREAS;			
RA	Zacchetti D., De Pietri Tonelli D., Schnurbus R.;			
RT	"New beta-site APP cleaving enzyme isoform (BACE-1C) obtained from			
RT	human pancreas.";			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB050437; BAB40932.1; -			
DR	EMBL; AF338817; AAK38375.1; -			
DR	HSSP; P32329; 1YPS.			
DR	InterPro: IPR001969; Asp. protease.			
DR	InterPro: IPR001461; Pepsin.			
DR	Pfam: PF00026; asp. 1.			
DR	PRINTS; PR00792; PEPSTIN.			
DR	PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.			
DR	SEQUENCE 457 AA; 51068 MW; C794C9A9E85FE7A2 CRC64;			
QY	Query Match	86.5%;	Score 1649;	DB 4; Length 457;
QY	Best Local Similarity	87.8%;	Pred. No. 2.2e-136;	
QY	Matches 317; Conservativity	0;	Mismatches 0;	Indels 44; Gaps
QY	1	MDNLRSKSGGYVEMTWGSPQTLNIILVDTGSSNFVAGADPHFLHRYQROLSSYTR	60	
QY	63	MDNLRSKSGGYVEMTWGSPQTLNIILVDTGSSNFVAGADPHFLHRYQROLSSYTR	122	
QY	61	DLRKGVPYPTQGWESGLGTDLYSIPHGPNVYRANIAAITESDFFLINGSWWEILGL	120	
QY	123	DLRKGVPYPTQGWESGLGTDL-----	145	
QY	121	AYAIARPDPSLIEFFPSLVKOTHPNLFSLQCGAGFLNQGSEVLASVGGSMIGSIDH	180	
QY	146	-----PDPSLIEFFPSLVKOTHPNLFSLQCGAGFLNQGSEVLASVGGSMIGSIDH	198	

```
OY 181 SLVGTSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDKSIYDSGTTNLRPKVF 240
Db 199 SLVGTSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDKSIYDSGTTNLRPKVF 258
OY 241 EAAVKSIAASTTEKFPDGFGLGQOLVCWQAGTTPMNIFFVLSLYLMGEVYNOSFRITIL 300
Db 259 EAAVKSIAASTTEKFPDGFGLGQOLVCWQAGTTPMNIFFVLSLYLMGEVYNOSFRITIL 318
OY 301 POQYLRPEVEDVATSDODCKFAISOSSTGYVMGAVIMEGFYVFDRAKRRIQFAVSACHV 360
Db 319 POQYLRPEVEDVATSDODCKFAISOSSTGYVMGAVIMEGFYVFDRAKRRIQFAVSACHV 378
OY 361 H 361
Db 379 H 379
```

```
RESULT 4
O9BYB9 PRELIMINARY; PRT; 432 AA.
AC O9BYB9;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE BETA-SITE APP CLEAVING ENZYME I-432.
GN BACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H., Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production."
RL Neurosci. Lett. 307:9-12(2001).
EMBL: AB050438; BAB40933.1; -.
DR HSSP; P32329; 1YPS.
DR InterPro; IPR001969; Asp_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 432 AA; 48212 MW; 96FC81E6F0EED01B CRC64;
```

```
Query Match 78.9%; Score 1504.5; DB 4; Length 432;
Best Local Similarity 80.9%; Pred. No. 1e-123;
Matches 292; Conservative 0; Mismatches 0; Indels 69; Gaps 1;
```

```
OY 1 MVDNLRGKSGGCGYVEMVTPSPOTLNLIVDTGSSNFVGAAPHPLHRYRQRLSTYR 60
Db 63 MVDNLRGKSGGCGYVEMVTPSPOTLNLIVDTGSSNFVGAAPHPLHRYRQRLSTYR 122
OY 61 DLKRGVVPYTOGKMEGLGTDLVSIHPGPNVTVRANIAITLSDKFFINSMNEGILGL 120
Db 123 DLKRGVVPYTOGKMEGLGTDLVSIHPGPNVTVRANIAITLSDKFFINSMNEGILGL 145
OY 121 AVAELARDDSLPEFFDSLVKQTHVPLNLSLQCGAGFPLNQSEVLASVSGSMIIGIDH 180
Db 146 -----LCGAGFPLNQSEVLASVSGSMIIGIDH 173
OY 181 SLVGTSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDKSIYDSGTTNLRPKVF 240
Db 174 SLVGTSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDKSIYDSGTTNLRPKVF 233
OY 241 EAAVKSIAASTTEKFPDGFGLGQOLVCWQAGTTPMNIFFVLSLYLMGEVYNOSFRITIL 300
Db 234 EAAVKSIAASTTEKFPDGFGLGQOLVCWQAGTTPMNIFFVLSLYLMGEVYNOSFRITIL 293
OY 301 POQYLRPEVEDVATSDODCKFAISOSSTGYVMGAVIMEGFYVFDRAKRRIQFAVSACHV 360
```

```
Db 294 POQYLRPEVEDVATSDODCKFAISOSSTGYVMGAVIMEGFYVFDRAKRRIQFAVSACHV 353
OY 361 H 361
Db 354 H 354
```

```
RESULT 5
O9H2V8 PRELIMINARY; PRT; 439 AA.
AC O9H2V8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CDAL3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PHOCHROMOCYTOMA;
RA Li Y., Huang Q., Peng, Y., Song H., Yu Y., Xu S., Ren S., Chen Z.,
RA Han Z.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF212252; AAC41783.1; -.
DR HSSP; P00797; ZREN.
DR InterPro; IPR001969; Asp_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp; 3.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 439 AA; 46275 MW; 02EC0E050F11602 CRC64;
```

```
Query Match 58.0%; Score 1106; DB 4; Length 439;
Best Local Similarity 56.4%; Pred. No. 1e-89;
Matches 202; Conservative 60; Mismatches 92; Indels 4; Gaps 2;
```

```
OY 1 MVDNLRGKSGGCGYVEMVTPSPOTLNLIVDTGSSNFVGAAPHPLHRYRQRLSTYR 60
Db 1 MVDNLRGKSGGCGYVEMVTPSPOTLNLIVDTGSSNFVGAAPHPLHRYRQRLSTYR 60
OY 61 DLKRGVVPYTOGKMEGLGTDLVSIHPGPNVTVRANIAITLSDKFFINSMNEGILGL 120
Db 61 SKGPDVYVKTQSGWNGFVGEDLVITPKGNITSFLVNIAITFSENFPLPGIKWNGILGL 120
OY 121 AVAELARDDSLPEFFDSLVKQTHVPLNLSLQCGAGFPLNQSEVLASVSGSMIIGIDH 180
Db 121 AVAELARDDSLPEFFDSLVKQTHVPLNLSLQCGAGFPLNQSEVLASVSGSMIIGIDH 177
OY 181 SLVGTSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDKSIYDSGTTNLRPKVF 240
Db 178 SLVGTSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDKSIYDSGTTNLRPKVF 237
OY 241 EAAVKSIAASTTEKFPDGFGLGQOLVCWQAGTTPMNIFFVLSLYLMGEVYNOSFRITIL 300
Db 238 DAVEAVARASLIPERSDFFWGSQACWTNSPTSPSYFKISITLRDNSSRSRITIL 297
OY 301 POQYLRPEVEDVATSDODCKFAISOSSTGYVMGAVIMEGFYVFDRAKRRIQFAVSACHV 358
Db 298 POLYIOPMAGAGLNY-ECYRFGISPTNALVIGATVMEGFYVFDRAKRRIQFAVSACHV 354
```

```
RESULT 6
O9JL18 PRELIMINARY; PRT; 514 AA.
AC O9JL18;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ASPARTYL PROTEASE 1.
GN BACE2.
```

RESULT 7		
09C0U5		
ID 09C0U5	PRELIMINARY;	PRT: 266 AA.
AC 09C0U5;		
DT 01-JUN-2001 (TREMBLrel, 17, Created)		
DT 01-JUN-2001 (TREMBLrel, 17, Last sequence update)		
DT 01-DEC-2001 (TREMBLrel, 19, Last annotation update)		
DE BETA-SITE APP CLEAVING ENZYME (FRAGMENT).		
GN BACE.		
OS Mus musculus (Mouse).		
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX NCBI_TaxID=10090;		
RN [1]		
RN SEQUENCE FROM N.A.		
RC STRAIN=C57BL/6J; TISSUE=BRAIN;		
RC MEDLINE=21085660; PubMed=11217651;		
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,		
RA Aitawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,		
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,		
RA Kato K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		

```

RESULT 8
09NZL1
AC 09NZL1 PRELIMINARY; PRT; 396 AA.
ID 09NZL1:
DT 01-OCT-2000 (TREMblrel_15, Created)
DT 01-OCT-2000 (TREMblrel_15, Last sequence update)
DT 01-DEC-2001 (TREMblrel_15, Last annotation update)
DE ASPARTYL PROTEASE.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de la Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase.";
RL Cytogenet. Cell Genet. 89:177-184(2000).
DR EMBL; AF188277; AAF38386.1; -.
DR HSP; P00797; 2REN.
DR MEROPS; A01.041; -.
DR InterPro; IPR001969; ASP_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp; 2.
DR PRINTS; PR00797; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KM Protease.

```


RP SEQUENCE FROM N.A.
RA Accarino M., Fumagalli P., Taramelli R., Ottolenghi S.;
RT "Cloning of a gene from chromosome 21 Down Region encoding a potential
RT transmembrane protease."; to the EMBL/GenBank/DBJ databases.
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF051150; AAD45964.1; -
DR MEROPS: A01.041; -
DR InterPro: IPR001969; Asp.protease.
DR InterPro: IPR001461; pepsin.
DR Pfam: PF00026; asp. 2.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
FT NON TER
SQ SEQUENCE 255 AA; 28685 MW; 53DE317815996D63 CRC64;
Query Match 28.7%; Score 547.5; DB 11; Length 255;
Best Local Similarity 56.7%; Pred. No. 5.2e-40;
Matches 97; Conservative 31; Mismatches 42; Indels 1; Gaps 1;
QY 188 WYPTIRREYEVIVIVVEINGODLKMCKEYNDKSIYDSCGTTNLRPKVFEAAVAKSI 247
DB 1 WYPTIKKRYVQIETIKLEIGGNNLDCREYNADKAIYDSCGTTLRPKVFEAAVAV 60
QY 248 KAASSTKFPDGFWMGLVCMQAGTTPNIPVSYLMGEVTSRITLIPQGYLRP 307
DB 61 ARTSLPEFSDGFWTGAOLACTNSETPWAVPPKISLYRDENASKRITLIPQGYLRP 120
QY 308 VEDVAISODDCYKFAISOSSTGTVMGAVIMEGFYVDFRARRKIGFAVASC 358
DB 121 MMGAGENTY-ECIRFGISSSTNALVIGATYMEGFYVDFRARRKIGFAVASC 170
RESULT 12
QYKPK6 PRELIMINARY; PRT; 391 AA.
AC 09VKP6; PRELIMINARY; PRT; 391 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG17134 PROTEIN.
GN CG17134.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanaides P.G., Scherer S.E., Li P.W., Hoskins R.G., Galle J.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.T., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Doming M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flosser C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Foster C., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaiswal B.E., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulyov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhao X., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC EMBL: AE003630; AAF53016.1; -
CC HSSP: P00794; 4CMS.
DR MEROPS: A01.0PW; -
DR FLYBASE: FBgn0032304; CG17134.
DR InterPro: IPR001969; Asp.protease.
DR InterPro: IPR001461; pepsin.
DR Pfam: PF00026; asp. 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
DR ASpartyl protease; Hydrolyase.
KW SEQUENCE 391 AA; 42016 MW; 99A18E7131025E1B CRC64;
Query Match 17.4%; Score 332.5; DB 5; Length 391;
Best Local Similarity 29.1%; Pred. No. 7.5e-21;
Matches 105; Conservative 61; Mismatches 146; Indels 49; Gaps 13;
QY 3 DNLKSGSGGYVEVEMTVPQPTNLIVDTGSSNFAVGAAPHP-----LHRYQRLS 56
DB 66 ENLHSMNNEYGYVIAIGTPEQREFNILEFTGSAWLWPSASCPASNTACQHRNKYSSAS 125
QY 57 STYRDLKRGVYVPTQGWBEGLDVLSPHGNVVRANIAITSDKFFINGSWMEG 116
DB 126 SYTVANGGEFFLEVTGSLSGFLSNDIVTIA-GISTONOTGEALSEPQTTFVD-AFAG 183
QY 117 ILGLAVETIARPDSDLPEFPDILVKQTHVPN-LFSLQCGAGFPLNOSSEVLASGSKIT 175
DB 184 ILGLAFSAIA--VDGVTTPFDNMISQGLDEPVSIFYLKRQ-----TAVNGGELIL 233
QY 176 GGIDSLVYSGSLWTPPIRREMYEVIVIVVEINGODLKMCKEYNDKSIYDSCGTTNLR 235
DB 234 GGIDSLVYSGSLWTPPIRREMYEVIVIVVEINGODLKMCKEYNDKSIYDSCGTTNLR 287
QY 236 PKKVFPAVKSIIKAASSTKFPDGFWMGLGEOLV-CMQAGTTPNIPVSYLMGEVTSRIT 294
DB 288 PLAAVAKIRKQGLANDND-----GEAFVRCGRVSS-----LPKYNLNGITV---- 329
QY 295 FRITLIPQGYLRPVEDVATSDQCKY-FAISOSSSTGTVMGAVIMEGFYVDFRARRKIGF 353
DB 330 --FTLAPRDYI--VKWTONGQTCMSAFYMGSLFWILGDVIFKFTYVDPDKMERIGF 385
QY 354 A 354
DB 386 A 386
RESULT 13
QYKPK6 PRELIMINARY; PRT; 354 AA.
AC 09GXY7; PRELIMINARY; PRT; 354 AA.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HEME-BINDING ASPARTIC PROTEINASE (FRAGMENT).
OS Boophilus microplus (Cattle tick).
OC Eukaryota; Metazoa; Arthropoda; Arachnida; Acari;

CC Parasitiformes; Ixodidae; Ixodidae; Boophilus.
 OX NCBI_TaxID=6941;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PORTO ALEGRE; TISSUE=OVARY;
 RA Sordane M.H.F., Logullo C., Zingali R.B., Paiva-Silva G.O.,
 RA Juliano L., Oliveira P.L.;
 RT "A heme-binding aspartic proteinase from the eggs of the hard tick
 RT Boophilus microplus.";
 RL J. Biol. Chem. 0:0-0(2000).
 DR EMBL: AF286865; AAC00993.1; -
 DR HSSP: P00797; 2REN.
 DR MEROPS: A01.054;
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp. 1.
 DR PRINTS: PR00792; PEPsin.
 FT NON_TER 1 1
 SQ SEQUENCE 354 AA; 38512 MW; B50D7C638CF27091 CRC64;
 Query Match 17.4%; Score 332; DB 5; Length 354;
 Best Local Similarity 26.2%; Pred. No. 7.2e-21;
 Matches 101; Conservative 74; Mismatches 135; Indels 76; Gaps 15;
 QY 2 VDNLRGK-----SGGGYVEMTVSGSPQTLNIVDTGSSNFAVGAAPHFL- 47
 DB 13 VTEIRGALDDPIPIITNTNNMQFIITIGTPQSFKLMDTSSNFW-----PSIN 67
 QY 48 -----HRYQROLSTYRDLRKGVVYPTQGMKEGLTDLVSIPIHGPNTVRANI 98
 DB 68 CDOSMACRDAKSDSSSTFTKSGRYIRIRYSGVVRGITSIDNGV--GPATVQYKF 125
 QY 99 AATSEKPFINGSNMGILGLAVAIARDDSLPEFPSLYKQTHVPN-LFSLQLGAG 157
 DB 126 AEMDHSDGKIFRNAKKGITGLFLAFPSISQ--NQDLPLFAMKQGVRAVPSLYL--SK 181
 QY 158 FPIQSEVLASVSGSMIGIDHSLTYGSLWYPIRRENVYEIVYRVINQODLKM-DC 216
 DB 182 QPEQN-----GGEIFYGGINAQRYTGAIHYVPSQAAMHQMVDNINVOGTTLCVGC 235
 QY 217 KEVNVKSTVDSTNTLRPKKVFEAAYKSIKAASTEKFPDGFWMGEQLVCMQAGTTPW 276
 DB 236 -----PTVVDSTSTFLSGP---SADVETLNKRVIGATKTAGY-----FEVNCATI 277
 QY 277 NIEPVISLYLME---VTNQSFRTITLPOQYLRPVEDVATSDDCYKFAISOSTGT--- 330
 DB 278 SSLEPITFLNKGKSPFLQGEATIRI-----PLTTGEGCFTRISESADGTMLW 327
 QY 331 WNGAVIMEGFVYVFDRAKRRIGFAVS 356
 DB 328 ILGAVFTQTYTVYFDRANQVRGFATA 353
 RESULT 14
 O9DEC2 PRELIMINARY; PRT; 384 AA.
 AC O9DEC2
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PEPsinogen A.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
 CC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ikuzawa M., Inokuchi T., Kobayashi K., Yasumasa S.;
 RT "Molecular Cloning of Pepsinogens in Adult Xenopus laevis and Bullfrog
 RT Rana catesbeiana.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.

DR EMBL: AB045380; BAB20798.1; -
 DR HSSP: P00790; 1PSN.
 DR MEROPS: A01.001; -
 DR InterPro: IPR001969; Asp-protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp. 1.
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE; 1.
 KW Aspartyl protease; Hydrolase.
 SQ SEQUENCE 384 AA; 41888 MW; 729DD2E7953D9072 CRC64;
 Query Match 16.2%; Score 308; DB 13; Length 384;
 Best Local Similarity 27.5%; Pred. No. 1e-18;
 Matches 99; Conservative 60; Mismatches 133; Indels 68; Gaps 16;
 QY 13 YVEMTVSGSPQTLNIVDTGSSNFAVGAAPHFL-----HRYQROLSTYRDLRK 65
 DB 72 YGTISIGTPPEFTVIDTGSANLW---PSYSSQACSNHNFNQOSTTQATNTP 128
 QY 66 VVYPTQGMKEGLTDLVSIPIHGPNTVRANIAITESDK-FPIGNSNMGILGLAVAE 124
 DB 129 VSIQGTGSMGFLGYDYLQV---GNIOISNQMGISESEGFYLYSPFDGILGLAFPS 185
 QY 135 IARPDSLEPFPSLYKQTHVP-NLFSLQAGPFLNQSEVLASVSGSMITGIDHSLY 183
 DB 186 IA--SSQATPVPDNNMWSGGLIPQNLFSYLLSDG-----QTGSYVLEGGVDSNY 233
 QY 184 TGSMLTYTIRRENVYEIVYRVINQODL--KMDCKEYNDKSTVDSTNTLRPKKVF 241
 DB 234 SGLSMVPLTAEYWQITLDSVINGVYIACSQC-----QALVDGTSLMTGSPPI- 286
 QY 242 AAVKSIKAASTEKFPDGFWMGEQLV-CMQAGTTPPNIPIVTSILYLMGEVYNQSFRTITL 300
 DB 287 ANIQVYIASQDSN-----GQYVINCNNISNPTTVF-----TIN 321
 QY 301 PQQY-LRPVEDVATSDDCYK-FAISOSTGT-----VMGAVIMEGFVYVFDRAKRRIGFA 354
 DB 322 GVQYPLSPSAVAVRONQCCSGFQAMNLPFTNSGDMILIGVFIHQTYTVDRANNVAYA 361
 RESULT 15
 O9DEC4 PRELIMINARY; PRT; 385 AA.
 AC O9DEC4
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PEPsinogen A.
 OS Rana catesbeiana (Bull frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
 OX NCBI_TaxID=8400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ikuzawa M., Inokuchi T., Kobayashi K., Yasumasa S.;
 RT "Molecular Cloning of Pepsinogens in Adult Xenopus laevis and Bullfrog
 RT Rana catesbeiana.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 DR EMBL: AB045376; BAB20092.1; -
 DR HSSP: P00794; 4CMS.
 DR MEROPS: A01.001; -
 DR InterPro: IPR001969; Asp-protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp. 1.
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE; 1.
 KW Aspartyl protease; Hydrolase.
 SQ SEQUENCE 385 AA; 41702 MW; 5DC8914FC9CEC603 CRC64;
 Query Match 16.2%; Score 308; DB 13; Length 385;
 Best Local Similarity 27.8%; Pred. No. 1e-18;

Matches 99; Conservative 64; Mismatches 133; Indels 60; Gaps 15;

```
QY 13 YVEMTVGSPQOTNLIVDTGSSNFAVG---AAPHPELHRYQROLSTYRDLRKGYV 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 YFGTISIGTPPQSFVIFDGTSSNLWVPSYCSSPACTNHMFNPQSSSTFQATNTPYSI 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 69 PYTOGKMEGELGDLVSIPHGPNVTVRANIAAITESDK-FFINGSNMEGILGLAYAEIAR 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 QYGTSMGSGFLGYDVQY---GNIOITNQIFGLSQSEPGSFLYXSPFDGILGLAFPSLA- 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 128 PDDSLPEPFDLSIVKOTHP-NLFSIQLCAGAPLNQSEVLASVGSMITIGIDHSLYTGS 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 -SSQATPVFDNMNMNOGLIPODLFSYLLSSQG---QS-----GSFVLEGGVDTSYTG 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 LMVTPIRREWYEVIIIVREINIGODLKM--DCKEYNDKSIYDSGTNLRPKVFEAAV 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 LMWVPLLAETVWQITVDISISIGQVIACSGSC-----SAIVDTGTSILAGP---STPI 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 245 KSIRKAASSTEKFPDGFWLGEQLVCMQAGTTPWNIFPVISLYIMGEVTNQSFRITILLPOQY 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 288 ANIQYVIGANDQSNQYV---INCNNISMPTVVF-----TINGVQY 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 305 LRPVED-VATSDDC---YKFAISQSTGT--VMGAVIMEGFYVEDRARRKRIGFA 354
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 327 PLPASAYVROSOQSTSGFOAMNLPYSSGDLMIIGDVFIREYVVPDRANNVYAMA 382
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: October 30, 2002, 12:30:03
Job time : 114.646 secs

6/15/98

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:20:50 ; Search time 144.577 Seconds
(without alignments)
277.344 Million cell updates/sec

Title: us-09-724-571-75

Perfect score: 1907
Sequence: 1 MVDMLRGKSGGYVEMTVG.....VFEDRRKRKIGFAVSACHVH 361

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1907	100.0	415	21	AA07899
2	1907	100.0	453	21	AA188438
3	1907	100.0	453	22	AAE10642
4	1907	100.0	453	22	AAE06872
5	1907	100.0	453	22	AAU06616
6	1907	100.0	453	22	AAU07215
7	1907	100.0	453	22	AAE02594
8	1907	100.0	456	21	AA07897
9	1907	100.0	459	21	AA188439
10	1907	100.0	459	22	AAE10643
11	1907	100.0	459	22	AAE06873

12	1907	100.0	459	22	AAU06617	Human-pro-Asp 2(a)
13	1907	100.0	459	22	AAU07216	Human aspartyl pro
14	1907	100.0	459	22	AAE02595	Human-Asp-2(a) del
15	1907	100.0	460	21	AA07898	Amino acid sequenc
16	1907	100.0	488	22	AA066572	Human memapsin 2.
17	1907	100.0	488	22	AA061334	Memapsin 2 protein
18	1907	100.0	501	21	AA194767	Human beta-secreta
19	1907	100.0	501	21	AA07896	Amino acid sequenc
20	1907	100.0	503	22	AA066573	Human pro-memapsin
21	1907	100.0	503	22	AA061335	T7 promoter and ve
22	1902	99.7	425	21	AA188437	Human Asp2 amino a
23	1902	99.7	425	22	AAE10641	T7-Caspase-Caspase
24	1902	99.7	425	22	AAE06871	T7-Caspase-Caspase
25	1902	99.7	425	22	AAU06615	T7-Caspase-Caspase
26	1902	99.7	425	22	AAU07214	T7-Caspase-Caspase
27	1902	99.7	425	22	AAE02593	Human-pro-Asp 2(a)
28	1902	99.7	433	21	AA188433	Human-pro-Asp 2(a)
29	1902	99.7	433	22	AAE10640	Human-pro-Asp 2(a)
30	1902	99.7	433	22	AAE06870	Human-pro-Asp 2(a)
31	1902	99.7	433	22	AAU06614	Human-pro-Asp 2(a)
32	1902	99.7	433	22	AAU07213	T7-human aspartyl
33	1902	99.7	433	22	AAE02592	Human-pro-Asp-2(a)
34	1902	99.7	446	21	AA188431	T7-caspase-human-P
35	1902	99.7	446	22	AAE10638	T7-Human-pro-Asp 2
36	1902	99.7	446	22	AAE06868	T7-Human-pro-Asp2
37	1902	99.7	446	22	AAU06612	Human T7-Human-pro
38	1902	99.7	446	22	AAU07211	T7-human aspartyl
39	1902	99.7	446	22	AAE02590	T7-Human-pro-Asp-2
40	1902	99.7	459	21	AA188432	T7-caspase-human-P
41	1902	99.7	459	22	AAE10639	T7-Caspase-human-P
42	1902	99.7	459	22	AAE06869	Human T7-Caspase-H
43	1902	99.7	459	22	AAU06613	T7-caspase-human a
44	1902	99.7	459	22	AAU07212	T7-caspase-human a
45	1902	99.7	459	22	AAE02591	T7-caspase-human-P

ALIGNMENTS

RESULT 1
AA07899 standard; Protein: 415 AA.
AC AA07899;
XX
XX
DT 14-NOV-2000 (first entry)
XX
XX
DE Amino acid sequence of a human beta-secretase enzyme fragment.
XX
KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor.
XX
XX Homo sapiens.
OS
XX
XX WO200047618-A2.
XX
XX PD 17-AUG-2000.
XX
XX PF 10-FEB-2000; 2000WO-US03819.
XX
XX PR 10-FEB-1999; 99US-0119571.
XX
XX PR 15-JUN-1999; 99US-0139172.
XX
XX PA (ELAN-) ELAN PHARM INC.
XX
XX PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
XX
XX Slnha S, Tatsuno G, Tung J, Wang S, McConlogue LJ,
XX
XX WPI: 2000-533011/48.
XX
XX Purified beta-secretase protein used in assays to discover inhibitors
XX
XX which can be used for the treatment of amyloidogenic diseases e.g.

PT Alzheimer's disease -
 XX Claim 10; Fig 3B; 121pp; English.
 PS
 XX The specification describes a beta-secretase enzyme. The enzyme cleaves
 CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
 CC enzyme is therefore implicated in the production of amyloid plaque
 CC components which accumulate in the brains of individuals afflicted with
 CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
 CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
 CC disease-like pathology to test if they maintain or improve cognitive
 CC ability or reduce the plaque burden. The compounds are used for the
 CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
 CC present sequence represents a human beta-secretase enzyme fragment.
 XX
 XX Sequence 415 AA:
 S0
 Query Match 100.0%; Score 1907; DB 21; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1.2e-191;
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDNLRGSGGQYVEMTVGSPPTLNLIVDTGSSNFVGAAPHFLHRRYOROLSTYR 60
 DB 18 MYDNLRGSGGQYVEMTVGSPPTLNLIVDTGSSNFVGAAPHFLHRRYOROLSTYR 77
 QY 61 DLKRGVYPTQGWKEGELGTDVSIIPHGPNTVRANTAAITSDKFFLNGSNMEGILGL 120
 DB 78 DLKRGVYPTQGWKEGELGTDVSIIPHGPNTVRANTAAITSDKFFLNGSNMEGILGL 137
 QY 121 AYAELIARPDSDLPEPFDLSLVKOTHPNLFSLDLCAGAPFLNOSSEVLASVGSMTIIGIDH 180
 DB 138 AYAELIARPDSDLPEPFDLSLVKOTHPNLFSLDLCAGAPFLNOSSEVLASVGSMTIIGIDH 197
 QY 181 SLYTGSLWTPPIRREMYEVIIVRVEINGODLKMCKEYNYKSIYDSGTTMLRLPKKVF 240
 DB 198 SLYTGSLWTPPIRREMYEVIIVRVEINGODLKMCKEYNYKSIYDSGTTMLRLPKKVF 257
 QY 241 EAAVKSIRKAASSTERFPDGFMLGEOLVCMQAQGTTPMNIIPVISLYLMGEVTNOSFRITIL 300
 DB 256 EAAVKSIRKAASSTERFPDGFMLGEOLVCMQAQGTTPMNIIPVISLYLMGEVTNOSFRITIL 317
 QY 301 POQYLRPVEDVATSDDDCKKFAISOSSTGTVMGAYIMEFYVVPDARRRIGFAVSACHV 360
 DB 318 POQYLRPVEDVATSDDDCKKFAISOSSTGTVMGAYIMEFYVVPDARRRIGFAVSACHV 377
 QY 361 H 361
 DB 378 H 378

RESULT 2
 AAY88438
 ID AAY88438 standard; Protein; 453 AA.
 XX
 AC AAY88438;
 XX
 DF 03-AUG-2000 (first entry)
 XX
 DE Modified human aspartyl protease 2 (Asp2) amino acid sequence.
 KW Aspartyl protease; aspartase; amyloid precursor protein; APP; Asp 2;
 KW Alzheimer's disease; beta secretase site.
 XX
 OS Homo sapiens.
 OS
 PN MO200017369-A2.
 XX
 PD 30-MAR-2000.
 XX
 PF 23-SEP-1999; 99WO-US20881.
 XX
 PR 24-SEP-1998; 98US-0101594.
 XX

PA (PHAA) PHARMACIA & UPJOHN CO.
 XX Gurney ME, Bienkowski MJ, Heinrikson RL, Parodi LA, Yan R;
 PI WPI: 2000-303209/26.
 XX N-PSDB; AAA15688.
 DR
 XX New enzyme designated human aspartase useful in research into
 PT Alzheimer's Disease is capable of cleaving amyloid protein precursor at
 PT the beta secretase site to produce amyloid beta peptide -
 XX
 XX Example 10; Page 169-172; 183pp; English.
 PS
 XX This sequence represents a modified human aspartyl protease 2 (Asp2)
 CC amino acid sequence. Asp2 encoded by this sequence has the C-terminal
 CC transmembrane domain deleted. The invention relates to a protease
 CC (e.g. Asp2) capable of cleaving the beta secretase site of amyloid
 CC precursor protein (APP). The protease contains a sequence encoding the
 CC amino acid sequence DTG and a sequence encoding DSG or DTG separated by
 CC 100-300 amino acids. When mutated the APP gene causes an autosomal
 CC dominant form of Alzheimer's disease. APP localises to the cell surface
 CC membrane and have a single C-terminal transmembrane domain. Proteolytic
 CC processing of APP produces the amyloid beta protein, which is possibly
 CC very important in Alzheimer's disease. The invention includes a
 CC nucleotide sequence encoding the protease, a vector containing the
 CC nucleotide sequence, and a cell line comprising the vector. Methods for
 CC screening for inhibitors of beta secretase activity are also given in the
 CC invention. The human aspartase protein and nucleotide sequences and the
 CC methods for identifying inhibitors of the protease, are useful in the
 CC treatment of and research in to Alzheimer's disease.
 XX
 S0 Sequence 453 AA:
 Query Match 100.0%; Score 1907; DB 21; Length 453;
 Best Local Similarity 100.0%; Pred. No. 1.3e-191;
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDNLRGSGGQYVEMTVGSPPTLNLIVDTGSSNFVGAAPHFLHRRYOROLSTYR 60
 DB 63 MYDNLRGSGGQYVEMTVGSPPTLNLIVDTGSSNFVGAAPHFLHRRYOROLSTYR 122
 QY 61 DLKRGVYPTQGWKEGELGTDVSIIPHGPNTVRANTAAITSDKFFLNGSNMEGILGL 120
 DB 123 DLKRGVYPTQGWKEGELGTDVSIIPHGPNTVRANTAAITSDKFFLNGSNMEGILGL 182
 QY 121 AYAELIARPDSDLPEPFDLSLVKOTHPNLFSLDLCAGAPFLNOSSEVLASVGSMTIIGIDH 180
 DB 183 AYAELIARPDSDLPEPFDLSLVKOTHPNLFSLDLCAGAPFLNOSSEVLASVGSMTIIGIDH 242
 QY 181 SLYTGSLWTPPIRREMYEVIIVRVEINGODLKMCKEYNYKSIYDSGTTMLRLPKKVF 240
 DB 243 SLYTGSLWTPPIRREMYEVIIVRVEINGODLKMCKEYNYKSIYDSGTTMLRLPKKVF 302
 QY 241 EAAVKSIRKAASSTERFPDGFMLGEOLVCMQAQGTTPMNIIPVISLYLMGEVTNOSFRITIL 300
 DB 303 EAAVKSIRKAASSTERFPDGFMLGEOLVCMQAQGTTPMNIIPVISLYLMGEVTNOSFRITIL 362
 QY 301 POQYLRPVEDVATSDDDCKKFAISOSSTGTVMGAYIMEFYVVPDARRRIGFAVSACHV 360
 DB 363 POQYLRPVEDVATSDDDCKKFAISOSSTGTVMGAYIMEFYVVPDARRRIGFAVSACHV 422
 QY 361 H 361
 DB 423 H 423

RESULT 3
 AAEL10642
 ID AAEL10642 standard; Protein; 453 AA.
 XX
 AC AAEL10642;
 XX
 DT 10-DEC-2001 (first entry)

XX	Human-Asp 2(a) protein lacking transmembrane domain.
DE	
XX	
KW	Human: aspartyl protease 2a; Asp2a; amyloid precursor protein; APP;
KM	Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
XX	amyloid plaque; neuronal loss; proteolytic; nototropic; neuroprotective.
KX	
OS	Homo sapiens.
XX	Synthetic.
FH	Key
FT	Misc-difference 214
XX	/note="Encoded by CAC"
PN	GB2357767-A.
XX	
PD	04-JUL-2001.
XX	
PF	22-SEP-2000; 2000GB-0023315.
XX	
XX	23-SEP-1999; 99US-0155493.
PR	23-SEP-1999; 99US-0404133.
PR	23-SEP-1999; 99WO-020881.
PR	13-OCT-1999; 99US-0416901.
PR	06-DEC-1999; 99US-0169232.
XX	
PA	(PHAA) PHARMACIA & UPJOHN CO.
PI	Bienkowski MJ, Gurney M;
DR	WPI: 2001-444208/48.
DR	N-PSDB: AAD17878.
XX	
XX	Polypeptide comprising fragments of human aspartyl protease with
PT	amyloid precursor protein processing activity and alpha-secretase
PT	activity, for identifying modulators useful in treating Alzheimer's
XX	disease.
XX	
PS	Example 10; Fig 11; 187pp; English.
XX	
CC	The patent discloses human aspartyl protease 1 (hu-Asp1) or modified
CC	Asp1 proteins which lack transmembrane domain or amino terminal
CC	domain or cytoplasmic domain and retains alpha-secretase activity
CC	and amyloid protein precursor (APP) processing activity. The proteins
CC	of the invention are useful for assaying hu-Asp1 alpha-secretase
CC	activity, which in turn is useful for identifying modulators of
CC	hu-Asp1 alpha-secretase activity, where modulators that increase
CC	hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's
CC	disease (AD) which causes progressive dementia with consequent
CC	formation of amyloid plaques, neurofibrillary tangles, gliosis and
CC	neuronal loss. hu-Asp1 protease substrate is useful for assaying
CC	hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with
CC	the substrate under acidic conditions and determining the level of
CC	hu-Asp1 proteolytic activity. The present sequence is human Asp 2(a)
CC	protein lacking a transmembrane (TM) domain. This sequence is generated
CC	by the deletion of the C-terminal TM domain of human Asp 2(a) protein.
XX	
S0	Sequence 453 AA;
XX	
Query Match	100.0%; Score 1907; DB 22; Length 453;
Best Local Similarity	100.0%; Pred. No. 1.3e-191;
Matches 361; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
0;	
0Y	1 MYDNLRGSGGGYEMVGVSPQTLNLIVDGGSNFVGAAPHFLHRYVROLSTYR 60
Db	63 MYDNLRGSGGGYEMVGVSPQTLNLIVDGGSNFVGAAPHFLHRYVROLSTYR 122
0Y	61 DLKRGVYPTTQGGKEGELGTDLVISPHGPNVTVRANIAATTESDKFFINGSNMGIIGL 120
Db	123 DLKRGVYPTTQGGKEGELGTDLVISPHGPNVTVRANIAATTESDKFFINGSNMGIIGL 182
0Y	121 AYAELARDDSLPEPFDLSLVQTHPNLFSQLCGAGFPLNDSKYLAVSGSMITGGIDH 180

Db	183	AMEIARERDDSLPEPFSDLVNQTHVNLFSIQLOGACFPPLN0SEVTLASVGGSGMTIGDII	242
Qy	181	SLYTSGLMYTPPIRREMYEVIIVREINGDOLKMDCKEYNDKSIYDSGTTNLRPKYF	240
Db	243	SLYTSGLMYTPPIRREMYEVIIVREINGDOLKMDCKEYNDKSIYDSGTTNLRPKYF	302
Qy	241	EAAVYSIRAASTKEPFGFWLGSLVQWAGTTPMNIIPVYISLYLMGEVTNDSFRITLL	300
Db	303	EAAVYSIRAASTKEPFGFWLGSLVQWAGTTPMNIIPVYISLYLMGEVTNDSFRITLL	362
Qy	301	POQYLRPEVDYATSDDDCYKFAISQSSTGYMGAVIMEGFYVDFDARRKIGFAVSACHV	360
Db	363	POQYLRPEVDYATSDDDCYKFAISQSSTGYMGAVIMEGFYVDFDARRKIGFAVSACHV	422
Qy	361	H 361	
Db	423	H 423	
RESULT 4			
AAE06872			
ID	AAE06872	standard; Protein; 453 AA.	
XX	AAE06872;		
AC			
XX			
DT	23-OCT-2001	(first entry)	
XX			
DE	Human-Asp2(a) deltatM protein.		
KM	Human; aspartyl protease 2a; Asp 2a; beta-amyloid precursor protein; APP		
KM	beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;		
KM	neurofibrillary tangle; neuronal loss; amyloid-beta peptide; neurotropic;		
KM	neuroprotective; antisense therapy; Asp2(a) deltatM protein;		
XX	gene therapy.		
OS	Homo sapiens.		
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	Misc-difference 214		
FT	/note= "Encoded by CAC"		
XX			
PN	WO200150829-A2.		
PD	19-JUL-2001.		
XX			
PF	09-MAY-2001; 2001WO-IB00799.		
XX			
PR	09-MAY-2001; 2001WO-IB00799.		
XX			
PA	(BIEN/) BIENKOWSKI M J.		
PA	(GURN/) GURNEY M E.		
PA	(HEIN/) HEINRIKSON R L.		
PA	(PARO/) PARODI L A.		
XX	(YANR/) YAN R.		
XX			
PI	Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;		
XX			
DR	WPI: 2001-483072/52.		
DR	N-PSDB: AADI3034.		
XX			
PT	Novel purified polypeptide comprising fragment of mammalian aspartyl		
PT	protease 2, lacking Asp2 transmembrane domain and retaining beta		
PT	secretase activity of Asp2 useful for identifying inhibitors of Asp2		
XX	activity		
XX			
PS	Claim 149; Fig 11; 185pp; English.		
XX			

therapeutics for the treatment and prevention of Alzheimer's disease, dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis and neuronal loss. App isoforms are also used in methods for identifying inhibitors and modulators of human Asp2 activity. The invention relates to a method for identifying agents that modulate the activity of human aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used as a means to screen in cellular assays for the inhibitors of beta- and gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in polymerase chain reactions (PCR). The probes are useful for detecting Hu-Asp nucleic acids in in vitro assays and in Northern and Southern blots. The present sequence is Human aspartyl protease 2a (Hu-Asp2a) deltaTM protein which is obtained by the deletion of transmembrane domain at the C-terminal end of Hu-Asp2a. Human Asp2a has beta-secretase activity.

Sequence 453 AA:

Query Match 100.0%; Score 1907; DB 22; Length 453;
 Best Local Similarity 100.0%; Pred. No. 1.3e-191;
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDNLRGKSGGGYVEMTVGSPPTNLILVDTGSSNFAGAAPHPFLHRYQRLSSTYR 60
 |||||||
 DB 63 MVDNLRGKSGGGYVEMTVGSPPTNLILVDTGSSNFAGAAPHPFLHRYQRLSSTYR 122
 |||||||
 QY 61 DLKRGVYVPTGKWEGLGTDLVSIPIHGPNVTVRANIAAITESDKFFINGSMWEGILGL 120
 |||||||
 DB 123 DLKRGVYVPTGKWEGLGTDLVSIPIHGPNVTVRANIAAITESDKFFINGSMWEGILGL 182
 |||||||
 QY 121 AYAEIARPPDSLPPFDSLTKQTHVNLFSLOLCGAGFPLNOSEVLASVGGSMIIGIDH 180
 |||||||
 DB 183 AYAEIARPPDSLPPFDSLTKQTHVNLFSLOLCGAGFPLNOSEVLASVGGSMIIGIDH 242
 |||||||
 QY 181 SLVTGSLMTYPIRREMYEVIIVRVEINQDILKMDCKEYNYKSIYDSGTTLRLPKKYF 240
 |||||||
 DB 243 SLVTGSLMTYPIRREMYEVIIVRVEINQDILKMDCKEYNYKSIYDSGTTLRLPKKYF 302
 |||||||
 QY 241 EAAVSIKKAASSTKRPDGFMLGEOLVCMQAGTTPWNIPVLSILMGSEVTSQSFRTITL 300
 |||||||
 DB 303 EAAVSIKKAASSTKRPDGFMLGEOLVCMQAGTTPWNIPVLSILMGSEVTSQSFRTITL 362
 |||||||
 QY 301 POQYLKRPEDVATSQDDCYKFAISQSGSTGTVMGAVIMEGFYVDFRARRKIGFAYSACHV 360
 |||||||
 DB 363 POQYLKRPEDVATSQDDCYKFAISQSGSTGTVMGAVIMEGFYVDFRARRKIGFAYSACHV 422
 |||||||
 QY 361 H 361
 ||
 DB 423 H 423
 ||

RESULT 5
 AAU06616
 ID AAU06616 standard; Protein; 453 AA.
 XX
 AC AAU06616;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE Human-pro-Asp 2(a) delta TM.
 XX
 KW Human; Aspartyl protease; beta-secretase; noctropic; ASP2;
 KM neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
 KW amyloid-beta; Abeta; Human-pro-Asp 2(a) delta TM; mutant; mutein.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Signal_peptide
 FT Protein 22..453
 FT /label= Mature_Human_pro-Asp_2(a)_delta_TM
 FT Misc-difference 214

/note="Encoded by CAC"

FT
 XX
 PN WC0200149098-A2.
 XX
 PD 12-JUL-2001.
 XX
 PF 09-MAY-2001; 2001WO-1B00798.
 XX
 PR 09-MAY-2001; 2001WO-1B00798.
 XX
 PA (BIEN/) BIENKOWSKI M J.
 PA (GURNEY/) GURNEY M E.
 PA (HEIN/) HEINRIKSON R L.
 PA (PARO/) PARODI L A.
 PA (YANR/) YAN R.
 PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;
 DR WPI; 2001-502549/55.
 DR N-PSDB; AAS11530.
 XX
 PT Novel purified polypeptide comprising fragment of mammalian aspartyl
 PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 PT activity
 PS Claim 149; Page 160; 185pp; English.
 XX
 CC The invention relates to a purified polypeptide comprising a fragment of
 CC mammalian aspartyl protease (Asp2) protein which lacks the Asp2
 CC transmembrane domain and the Asp2 protein, and where the polypeptide and
 CC the fragment retain the beta-secretase activity of the mammalian Asp2
 CC protein. The invention also details polynucleotides for the Asp
 CC proteins and vectors expressing them, and a polypeptide (isoform of
 CC amyloid protein precursor (APP)) comprising the amino acid sequence of an
 CC APP or its fragment containing an APP cleavage site recognizable by a
 CC mammalian beta-secretase, and further comprising two lysine residues at
 CC the carboxyl terminus of the amino acid sequence of the mammalian APP or
 CC APP fragment. Also included in the invention are methods of identifying
 CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
 CC useful for treating Alzheimer's disease. APP is useful in methods for
 CC identifying inhibitors or modulators of human Asp2 activity and
 CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
 CC therapeutics for the treatment or prevention of Alzheimer's disease.
 CC APP comprising the APP-sw beta-secretase peptide sequence (NMDA), which
 CC is associated with increased levels of Abeta processing is useful in
 CC assays relating to Alzheimer's research. The expression vector is useful
 CC for recombinantly expressing APP. Nucleic acids that hybridise to
 CC APP oligonucleotides are useful as probes or primers. The probes are
 CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
 CC Northern and Southern blots. The present sequence is Human-pro-
 CC Asp 2(a) delta TM protein, which lacks the C-terminal transmembrane
 CC domain.
 XX
 XX
 Sequence 453 AA:

Query Match 100.0%; Score 1907; DB 22; Length 453;
 Best Local Similarity 100.0%; Pred. No. 1.3e-191;
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDNLRGKSGGGYVEMTVGSPPTNLILVDTGSSNFAGAAPHPFLHRYQRLSSTYR 60
 |||||||
 DB 63 MVDNLRGKSGGGYVEMTVGSPPTNLILVDTGSSNFAGAAPHPFLHRYQRLSSTYR 122
 |||||||
 QY 61 DLKRGVYVPTGKWEGLGTDLVSIPIHGPNVTVRANIAAITESDKFFINGSMWEGILGL 120
 |||||||
 DB 123 DLKRGVYVPTGKWEGLGTDLVSIPIHGPNVTVRANIAAITESDKFFINGSMWEGILGL 182
 |||||||
 QY 121 AYAEIARPPDSLPPFDSLTKQTHVNLFSLOLCGAGFPLNOSEVLASVGGSMIIGIDH 180
 |||||||
 DB 183 AYAEIARPPDSLPPFDSLTKQTHVNLFSLOLCGAGFPLNOSEVLASVGGSMIIGIDH 242
 |||||||
 QY 181 SLVTGSLMTYPIRREMYEVIIVRVEINQDILKMDCKEYNYKSIYDSGTTLRLPKKYF 240
 |||||||

Db 243 SLVTSGLWYTPRRREWEYEVIIIVREINGODLKMCKEYNDKSIYDSCGTNLRPKKVF 302
 QY 241 EAAVKSIRKASSTKEKFPDGFMLGEOVLCMOAGTTPMWNIFPVISLYLMEGVNOSFRITIL 300
 Db 303 EAAVKSIRKASSTKEKFPDGFMLGEOVLCMOAGTTPMWNIFPVISLYLMEGVNOSFRITIL 362
 QY 301 POQYLRPEVDVATSDDCYKFAISQSTGTVMGAVIMEGFYVFDRAKRRIKGFVNSACHV 360
 Db 363 POQYLRPEVDVATSDDCYKFAISQSTGTVMGAVIMEGFYVFDRAKRRIKGFVNSACHV 422
 QY 361 H 361
 Db 423 H 423

RESULT 6
 AAU07215
 ID AAU07215 standard; Protein: 453 AA.
 AC AAU07215;

24-OCT-2001 (first entry)

Human aspartyl protease 2a deltatm (HuAsp-2adeltatm).

KM Human: aspartyl protease 1; Asp-1; neotropic; neuroprotective;
 KM aspartyl protease 2; Asp2; amyloid protein precursor; APP;
 KM beta-secretase; Alzheimer's disease; HuAsp-2adeltatm.

OS Homo sapiens.

Key Location/Qualifiers

FT Peptide 1..21

FT Protein /note="Signal peptide"

FT /note="Mature human aspartyl protease 2a deltatm"

PN MO200149097-A2.

PD 12-JUL-2001.

PF 09-MAY-2001; 2001MO-1B00797.

PR 09-MAY-2001; 2001MO-1B00797.

PA (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

PA (YANR/) YAN R.

PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;

DR WPI: 2001-502548/55.

DR N-PSDB: AAS11715.

Novel purified polypeptide comprising fragment of mammalian aspartyl
 protease 2, lacking Asp2 transmembrane domain and retaining beta
 secretase activity of Asp2 useful for identifying inhibitors of Asp2
 activity

Claim 149; Fig 11; 185pp; English.

The invention relates to a novel purified polypeptide comprising a
 fragment of mammalian aspartyl protease 2 (Asp2) protein which lacks the
 Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide
 and the fragment retain the beta-secretase activity of the mammalian Asp2
 protein. Also included is an isoform of amyloid protein precursor (APP)
 comprising the amino acid sequence of a APP or its fragment containing
 an APP cleavage site recognizable by a mammalian beta-secretase, and
 further comprising two lysine residues at the carboxyl terminus of the
 amino acid sequence of the mammalian APP or APP fragment. The

CC polypeptides are used for assaying for modulators of beta-secretase
 CC activity; identifying agents that inhibit the APP processing activity
 CC of human Asp2 aspartyl protease (Hu-Asp2); identifying agents that
 CC modulate the activity of Asp2; and for reducing cellular production of
 CC amyloid beta (Abeta) from APP. Agents identified by the above methods
 CC are useful for treating Alzheimer's disease, and for identifying
 CC modulators of amyloid-beta (Abeta) peptide production, for use in
 CC designing therapeutics for the treatment or prevention of Alzheimer's
 CC disease. Probes and primers derived from Asp nucleic acid sequences
 CC are useful for detecting Hu-Asp nucleic acids in vitro assays and in
 CC Northern and Southern blots. The present sequence represents the amino
 CC acid sequence of human Asp-2a delta TM construct which lacks the amino
 CC transmembrane domain. This construct was used for bacterial expression
 CC and purification of human Asp2a.

Sequence 453 AA;

Query Match 100.0%; Score 1907; DB 22; Length 453;
 Best Local Similarity 100.0%; Pred. No. 1.3e-191;
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDNLRKSGGQGYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYOROLSTYR 60
 Db 63 MDNLRKSGGQGYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYOROLSTYR 122
 QY 61 DLKRGYVYPTGKMGELGTDLVSTPHGPNTVVRANIAITESDKFTNGSNMEGIIGL 120
 Db 123 DLKRGYVYPTGKMGELGTDLVSTPHGPNTVVRANIAITESDKFTNGSNMEGIIGL 182
 QY 121 AYAELARPDSLEPPFDSLVKQTHVNFSLQCGAGPFLNOSVLAAGSGSMITGIDH 180
 Db 183 AYAELARPDSLEPPFDSLVKQTHVNFSLQCGAGPFLNOSVLAAGSGSMITGIDH 242
 QY 181 SIXTGLWYTPRRREWEYEVIIIVREINGODLKMCKEYNDKSIYDSCGTNLRPKKVF 240
 Db 243 SIXTGLWYTPRRREWEYEVIIIVREINGODLKMCKEYNDKSIYDSCGTNLRPKKVF 302
 QY 241 EAAVKSIRKASSTKEKFPDGFMLGEOVLCMOAGTTPMWNIFPVISLYLMEGVNOSFRITIL 300
 Db 303 EAAVKSIRKASSTKEKFPDGFMLGEOVLCMOAGTTPMWNIFPVISLYLMEGVNOSFRITIL 362
 QY 301 POQYLRPEVDVATSDDCYKFAISQSTGTVMGAVIMEGFYVFDRAKRRIKGFVNSACHV 360
 Db 363 POQYLRPEVDVATSDDCYKFAISQSTGTVMGAVIMEGFYVFDRAKRRIKGFVNSACHV 422
 QY 361 H 361
 Db 423 H 423

RESULT 7
 AAE02594
 ID AAE02594 standard; Protein: 453 AA.
 AC AAE02594;

10-AUG-2001 (first entry)

Human-Asp-2(a) delta TM protein.

Human: alpha-secretase; amyloid precursor protein; APP; therapy;
 KM Alzheimer's disease; antialzheimer's; aspartyl protease 2a; Asp2a;
 KM beta-secretase; Asp-2a delta TM.

OS Homo sapiens.

OS Synthetic.

Key Location/Qualifiers
 FT Msc-difference 214
 FT /note="Encoded by CAC"

PN MO200123533-A2.

```

PD 05-APR-2001.
XX
PF 22-SEP-2000; 2000MO-US26080.
XX
PR 23-SEP-1999; 99US-0155493.
XX 23-SEP-1999; 99MO-US20881.
PR 13-OCT-1999; 99US-0416901.
XX 06-DEC-1999; 99US-0169232.
XX
PA (PHAA ) PHARMACIA 6 UPJOHN CO.
PI Gurney M, Bienkowski MJ;
XX
XX WPI; 2001-290516/30.
DR N-PSDB; AAD06752.
XX
XX Enzymes that cleave the alpha-secretase site of the amyloid precursor
PT protein, useful for the treatment of Alzheimer's disease -
XX
XX Example 10; Fig 11; 189pp; English.
XX
XX The present invention relates to enzymes for cleaving the alpha-
CC secretase site of the amyloid precursor protein (APP) and methods of
CC identifying those enzymes. The methods may be used to identify enzymes
CC that may be used to cleave the alpha-secretase cleavage site of the APP
CC protein. The enzymes may be used to treat or modulate the progress of
CC Alzheimer's disease. The present sequence is human Aspartyl protease 2a
CC (Asp-2a) deltam protein which is obtained by deleting its transmembrane
CC domain. This sequence has beta-secretase protease activity.
XX
XX Sequence 453 AA:
SQ
Query Match 100.0%; Score 1907; DB 22; Length 453;
Best Local Similarity 100.0%; Pred. No. 1.3e-191; Indels 0; Gaps 0;
Matches 361; Conservative 0; Mismatches 0;
QY 1 MDNLTGKSGGQGYVEMTVGSPPTLNIIVDTGSSNFAGAAPHPFLHRYGROLSSTYR 60
DB 63 MDNLTGKSGGQGYVEMTVGSPPTLNIIVDTGSSNFAGAAPHPFLHRYGROLSSTYR 122
QY 61 DLKRGVYVPYTGKMGEGELGTDLVSIIPHGPNTVVRANIAITSDKFFNGSNWEGILGL 120
DB 123 DLKRGVYVPYTGKMGEGELGTDLVSIIPHGPNTVVRANIAITSDKFFNGSNWEGILGL 182
QY 121 AYAEIARPPDLSLPPFDSLVKQTHVNLFSLOLCGAGFPLNOLSEVLASVGSMTIGIDH 180
DB 183 AYAEIARPPDLSLPPFDSLVKQTHVNLFSLOLCGAGFPLNOLSEVLASVGSMTIGIDH 242
QY 181 SLVTGSLMTYPIRREWEYEVIIIVREINGODLKMCKEYNYDKSIYDSGTTNLRPKYF 240
DB 243 SLVTGSLMTYPIRREWEYEVIIIVREINGODLKMCKEYNYDKSIYDSGTTNLRPKYF 302
QY 241 EAAVKSTIKAASTKPFDPGFWLGBOLVCWQAGTTPNNIIFVVISLYLMGEVTONSFRITIL 300
DB 303 EAAVKSTIKAASTKPFDPGFWLGBOLVCWQAGTTPNNIIFVVISLYLMGEVTONSFRITIL 362
QY 301 POOYLREVEDVATSDODCYFAISQSSTGTVMGAVIMEGFYVVDRAKRRIQFAVSACHV 360
DB 363 POOYLREVEDVATSDODCYFAISQSSTGTVMGAVIMEGFYVVDRAKRRIQFAVSACHV 422
QY 361 H 361
DB 423 H 423

```

```

DE Active enzyme portion of human beta-secretase enzyme.
XX
XX beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KM amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
XX inhibitor; SS.
XX
XX Homo sapiens.
OS
XX
XX WO200047618-A2.
PN
XX 17-AUG-2000.
PD
XX
XX 10-FEB-2000; 2000MO-US03819.
XX
XX 10-FEB-1999; 99US-0119571.
XX 15-JUN-1999; 99US-0139172.
PR
XX (ELAN-) ELAN PHARM INC.
XX
XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sluhs S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX
XX WPI; 2000-533011/48.
DR
XX
XX Claim 24; Fig 2B; 121pp; English.
XX
XX The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents the active enzyme portion of human
CC beta-secretase enzyme.
XX
XX Sequence 456 AA:
SQ
Query Match 100.0%; Score 1907; DB 21; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.3e-191; Indels 0; Gaps 0;
Matches 361; Conservative 0; Mismatches 0;
QY 1 MDNLTGKSGGQGYVEMTVGSPPTLNIIVDTGSSNFAGAAPHPFLHRYGROLSSTYR 60
DB 18 MDNLTGKSGGQGYVEMTVGSPPTLNIIVDTGSSNFAGAAPHPFLHRYGROLSSTYR 77
QY 61 DLKRGVYVPYTGKMGEGELGTDLVSIIPHGPNTVVRANIAITSDKFFNGSNWEGILGL 120
DB 78 DLKRGVYVPYTGKMGEGELGTDLVSIIPHGPNTVVRANIAITSDKFFNGSNWEGILGL 137
QY 121 AYAEIARPPDLSLPPFDSLVKQTHVNLFSLOLCGAGFPLNOLSEVLASVGSMTIGIDH 180
DB 138 AYAEIARPPDLSLPPFDSLVKQTHVNLFSLOLCGAGFPLNOLSEVLASVGSMTIGIDH 197
QY 181 SLVTGSLMTYPIRREWEYEVIIIVREINGODLKMCKEYNYDKSIYDSGTTNLRPKYF 240
DB 198 SLVTGSLMTYPIRREWEYEVIIIVREINGODLKMCKEYNYDKSIYDSGTTNLRPKYF 257
QY 241 EAAVKSTIKAASTKPFDPGFWLGBOLVCWQAGTTPNNIIFVVISLYLMGEVTONSFRITIL 300
DB 258 EAAVKSTIKAASTKPFDPGFWLGBOLVCWQAGTTPNNIIFVVISLYLMGEVTONSFRITIL 317
QY 301 POOYLREVEDVATSDODCYFAISQSSTGTVMGAVIMEGFYVVDRAKRRIQFAVSACHV 360
DB 318 POOYLREVEDVATSDODCYFAISQSSTGTVMGAVIMEGFYVVDRAKRRIQFAVSACHV 377
QY 361 H 361

```

Db 378 H 378

RESULT 9
AAV88439
ID AAV88439 standard; Protein; 459 AA.
XX
AC AAV88439;
XX
DT 03-AUG-2000 (first entry)
XX
DE Modified human aspartyl protease 2 (Asp2) amino acid sequence.
XX
KW Aspartyl protease; aspartase; amyloid precursor protein; APP; Asp 2;
KW Alzheimer's disease; beta secretase site.
XX
OS Homo sapiens.
XX
PN MO200017369-A2.
XX
PD 30-MAR-2000.
XX
PF 23-SEP-1999; 99MO-US20881.
XX
PR 24-SEP-1998; 98US-0101594.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Gurney ME, Bienkowski MJ, Heintzlikson RL, Parodi LA, Yan R;
XX
DR WPI: 2000-303209/26.
XX
DR N-PSDB: AAA15689.
XX
PT New enzyme designated human aspartase useful in research into
PT Alzheimer's disease is capable of cleaving amyloid protein precursor at
PT the beta secretase site to produce amyloid beta peptide
XX
PS Example 10; Page 173-176; 183pp; English.

CC This sequence represents a modified human aspartyl protease 2 (Asp2)
CC amino acid sequence. Asp2 encoded by this sequence has the C-terminal
CC transmembrane domain deleted. The invention relates to a protease
CC (e.g. Asp2) capable of cleaving the beta secretase site of amyloid
CC precursor protein (APP). The protease contains a sequence separated by
CC amino acid sequence DTG and a sequence encoding DSG or DTG separated by
CC 100-300 amino acids. When mutated the APP gene causes an autosomal
CC dominant form of Alzheimer's disease. APP localises to the cell surface
CC membrane and have a single C-terminal transmembrane domain. Proteolytic
CC processing of APP produces the amyloid beta protein, which is possibly
CC very important in Alzheimer's disease. The invention includes a
CC nucleotide sequence encoding the protease, a vector containing the
CC screening for inhibitors of beta secretase activity are also given in the
CC invention. The human aspartase protein and nucleotide sequences and the
CC methods for identifying inhibitors of the protease, are useful in the
CC treatment of and research in to Alzheimer's disease.
XX
SQ Sequence 459 AA:

Query Match 100.0%; Score 1907; DB 21; Length 459;
Best Local Similarity 100.0%; Pred. No. 1,4e-191;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDNRGSGGCGYVEMVSGSPQTLNLVDTGSSNFVGAAPHEFLHRYQROLSTYR 60
DB 63 MYDNRGSGGCGYVEMVSGSPQTLNLVDTGSSNFVGAAPHEFLHRYQROLSTYR 122
QY 61 DLKRGVYPTQGWEGELGTDLVSIPIPGPNVYRANIAATTESKPFINSNMEGITGL 120
DB 123 DLKRGVYPTQGWEGELGTDLVSIPIPGPNVYRANIAATTESKPFINSNMEGITGL 182
QY 121 AYAEIARPDSDLPEPFDLSLVKQTHVNLFSLQCGAGFPLNQEVLASVSGSMITIGIDH 180

Db 183 AYAEIARPDSDLPEPFDLSLVKQTHVNLFSLQCGAGFPLNQEVLASVSGSMITIGIDH 242
QY 181 SLYTGSLWYTPIRREMYEVIIIVREINGODLKMDCKEYNKSIVDSGTTMLRIPKVF 240
DB 243 SLYTGSLWYTPIRREMYEVIIIVREINGODLKMDCKEYNKSIVDSGTTMLRIPKVF 302
QY 241 EAAVKSIKAASSTKPPDGFMLGEOIWCWQAGTTPWNIFFVLSYLMGEVTSFRITL 300
DB 303 EAAVKSIKAASSTKPPDGFMLGEOIWCWQAGTTPWNIFFVLSYLMGEVTSFRITL 362
QY 301 PQOYLRPVEDVATSDDDCKFAISOSTGTVMGAVIMEGFYVFPARRKRIGFANVACHV 360
DB 363 PQOYLRPVEDVATSDDDCKFAISOSTGTVMGAVIMEGFYVFPARRKRIGFANVACHV 422
QY 361 H 361
DB 423 H 423

RESULT 10
AAE10643
ID AAE10643 standard; Protein; 459 AA.
XX
AC AAE10643;
XX
DT 10-DEC-2001 (first entry)
XX
DE Human-Asp 2(a) protein with (His)6 tag and lacking TM domain.
XX
KW Human; aspartyl protease 2a; Asp2a; amyloid precursor protein; APP;
KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
KW amyloid plaque; neuronal loss; proteolytic; neurotropic; neuroprotective.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..21
FT Peptide
FT Protein
FT /label= Signal_peptide
FT /note= 22..459
FT MISC-difference 214
FT /note= "Mature Human-Asp2(a) delcATM (His)6 protein"
FT MISC-difference 454
FT /note= "Encoded by CAC"
FT MISC-difference 455
FT /note= "Encoded by CAG"
FT MISC-difference 456
FT /note= "Encoded by CAG"
FT MISC-difference 456
FT /note= "Encoded by CAG"
FT MISC-difference 457
FT /note= "Encoded by CAG"
FT MISC-difference 458
FT /note= "Encoded by CAG"
FT MISC-difference 459
FT /note= "Encoded by CAG"
FT MISC-difference 459
FT /note= "Encoded by CAG"
XX
XX GB2357767-A.
XX
XX
XX 04-JUL-2001.
XX
XX 22-SEP-2000; 2000GB-0023315.
XX
XX
XX 23-SEP-1999; 99US-0155493.
XX 23-SEP-1999; 99US-0404133.
XX 23-SEP-1999; 99MO-US20881.
XX 13-OCT-1999; 99US-0416901.
XX 06-DEC-1999; 99US-0169232.
XX
XX (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Bienkowski MJ, Gurney M;
XX
XX

DR WPI: 2001-444208/48.
 DR N-PSDB: AAD1879.
 XX Polypeptide comprising fragments of human aspartyl protease with
 PT amyloid precursor protein processing activity and alpha-secretase
 PT activity, for identifying modulators useful in treating Alzheimer's
 PT disease -
 XX
 PS Example 10; Fig 12; 187pp: English.
 XX The patent discloses human aspartyl protease 1 (hu-Asp1) or modified
 CC Asp1 proteins which lack transmembrane domain or amino terminal
 CC domain or cytoplasmic domain and retains alpha-secretase activity
 CC and amyloid protein precursor (APP) processing activity. The proteins
 CC of the invention are useful for assaying hu-Asp1 alpha-secretase
 CC activity, which in turn is useful for identifying modulators of
 CC hu-Asp1 alpha-secretase activity, where modulators that increase
 CC hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's
 CC disease (AD) which causes progressive dementia with consequent
 CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
 CC neuronal loss. Hu-Asp1 protease substrate is useful for assaying
 CC hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with
 CC the substrate under acidic conditions and determining the level of
 CC hu-Asp1 proteolytic activity. The present sequence is human Asp 2(a)
 CC protein lacking a transmembrane (TM) domain and containing (his)6
 CC sequence. This sequence is generated from human Asp 2(a) protein by
 CC the deletion of its C-terminal TM domain and addition of hexa-histidine
 CC tag at its C-terminus.
 CC
 XX Sequence 459 AA:
 SQ
 Query Match 100.0%; Score 1907; DB 22; Length 459;
 Best Local Similarity 100.0%; Pred. No. 1.4e-191;
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MPDNLKRGSGGQYVEMTVGSPPTLNIIVDTGSSNFAVGAAPHPLHRYGROLSSTYR 60
 63 MPDNLKRGSGGQYVEMTVGSPPTLNIIVDTGSSNFAVGAAPHPLHRYGROLSSTYR 122
 61 DLKRGVYVYTGQKMGELGDTLVSPHGPNTYVRANIAITSEDKFFINGSMWESILGL 120
 123 DLKRGVYVYTGQKMGELGDTLVSPHGPNTYVRANIAITSEDKFFINGSMWESILGL 182
 121 AVAEIARPDSDLEPPFDSLVKQTHVNLFSLOCGAGPFLNQEVLASVGGSMITIGIDH 180
 183 AVAEIARPDSDLEPPFDSLVKQTHVNLFSLOCGAGPFLNQEVLASVGGSMITIGIDH 242
 181 SLVTGSLMTYPIRREMYEVIIVVEINQDLKMDCKEYNDKSIYDSGTTNRLPKKYF 240
 243 SLVTGSLMTYPIRREMYEVIIVVEINQDLKMDCKEYNDKSIYDSGTTNRLPKKYF 302
 241 EAAVKSITKAASSTFEKPPDGMVLGEOLVCMQAGTPNNIFVILVLYMGEVTONOSFRITL 300
 303 EAAVKSITKAASSTFEKPPDGMVLGEOLVCMQAGTPNNIFVILVLYMGEVTONOSFRITL 362
 301 POOYLRFVEDVATFSDDCYCFATISQSGTGMGAVIMEGFYVVEDRARRKRGFAVSACHV 360
 363 POOYLRFVEDVATFSDDCYCFATISQSGTGMGAVIMEGFYVVEDRARRKRGFAVSACHV 422
 361 H 361
 423 H 423
 RESULT 11
 AAE06873
 ID AAE06873 standard; Protein; 459 AA.
 XX
 AC AAE06873;
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE Human-Asp2(a) deltaTM (His)6 protein.

XX Human: aspartyl protease 2a; Asp 2a; beta-amyloid precursor protein; APP;
 KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
 KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; neurotrophic;
 KW neuroprotective; antisense therapy; Asp2(a) deltaTM (His)6 protein;
 KW gene therapy.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Peptide 22..30
 FT /label= Signal_peptide
 FT Protein 31..459
 FT /note= "Mature Human-Asp2(a) deltaTM (His)6 protein"
 FT Region 1..22
 FT /note= "Corresponds to N-terminal Human-Asp2(a) deltaTM
 FT (His)6 protein"
 FT Misc-difference 214
 FT /note= "Encoded by CAC"
 FT Misc-difference 254
 FT /note= "Encoded by CAG"
 FT Misc-difference 255
 FT /note= "Encoded by CAG"
 FT Misc-difference 256
 FT /note= "Encoded by CAG"
 FT Misc-difference 257
 FT /note= "Encoded by CAG"
 FT Misc-difference 258
 FT /note= "Encoded by CAG"
 FT Misc-difference 259
 FT /note= "Encoded by CAG"
 FT Misc-difference 259
 FT /note= "Encoded by CAG"
 PD WO200150829-A2.
 XX 19-JUL-2001.
 PF 09-MAY-2001; 2001WO-IB00799.
 XX 09-MAY-2001; 2001WO-IB00799.
 PR 09-MAY-2001; 2001WO-IB00799.
 PA (BIEN/) BIENKOWSKI M J.
 PA (GURNE/) GURNEY M E.
 PA (HEIN/) HEINRIKSON R L.
 PA (PARO/) PARODI L A.
 PA (YANR/) YAN R.
 PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;
 XX WPI: 2001-483072/52.
 DR N-PSDB: AAD13035.
 PT Novel purified polypeptide comprising fragment of mammalian aspartyl
 PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of beta
 PT activity -
 XX
 PS Claim 149; Fig 12; 185pp: English.
 XX
 CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
 CC precursor protein (APP) isoforms and their corresponding DNA molecules.
 CC Human aspartyl proteases can act as beta-secretase proteases useful for
 CC treating Alzheimer's disease. APP isoforms are useful for identifying
 CC modulators of amyloid-beta peptide production, for use in designing
 CC therapeutics for the treatment and prevention of Alzheimer's disease,
 CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
 CC and neuronal loss. APP isoforms are also used in methods for identifying
 CC inhibitors and modulators of human Asp2 activity. The invention relates
 CC to a method for identifying agents that modulate the activity of human
 CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
 CC as a means to screen in cellular assays for the inhibitors or primers in
 CC gamma-secretase. Hu-Asp DNA fragments are useful as probes or primers in
 CC polymerase chain reactions (PCR). The probes are useful for detecting

CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
CC blots. The present sequence is Human aspartyl protease 2a (Hu-Asp2a)
CC delatm (His)6 protein which is obtained by deletion of C-terminal
CC transmembrane domain and addition of a hexa-Histidine tag at the
CC C-terminal end of Hu-Asp2a. Human Asp2a has beta-secretase activity.
XX

SO Sequence 459 AA:

Query Match 100.0%; Score 1907; DB 22; Length 459;
Best Local Similarity 100.0%; Pred. No. 1,4e-191;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDNLRGKSGGGYVEMTGSPPTNTLIVDTGSSNFVGAAPHPLHRYRQRLSTYR 60
DB 63 MYDNLRGKSGGGYVEMTGSPPTNTLIVDTGSSNFVGAAPHPLHRYRQRLSTYR 122
QY 61 DLKRGVYVYTGKMGELGTDLVSIPIGPNVTYRANIAATTESDKFINSNMEGITGL 120
DB 123 DLKRGVYVYTGKMGELGTDLVSIPIGPNVTYRANIAATTESDKFINSNMEGITGL 182
QY 121 AYAETARPDDSLPEPFDSLVKQTHVFNLFSLQCGAGFPLNQSEVLASVGSMTIGIDH 180
DB 183 AYAETARPDDSLPEPFDSLVKQTHVFNLFSLQCGAGFPLNQSEVLASVGSMTIGIDH 242
QY 181 SLYTGSIMWTPIRREMYEVIIVREINQDLMCKEYNYDKSIYDSGTTLRLPKKVF 240
DB 243 SLYTGSIMWTPIRREMYEVIIVREINQDLMCKEYNYDKSIYDSGTTLRLPKKVF 302
QY 241 EAAVSKIAASSTEFKPDGFWLGEOLVQWAGTTPMNIIPVLSILMGVTVNQSFRITL 300
DB 303 EAAVSKIAASSTEFKPDGFWLGEOLVQWAGTTPMNIIPVLSILMGVTVNQSFRITL 362
QY 301 POQYLRPEVDVATSDDDCKFAISQSSSTGTVMGAVIMEGFYVFPARRKRIGFAVSACHV 360
DB 363 POQYLRPEVDVATSDDDCKFAISQSSSTGTVMGAVIMEGFYVFPARRKRIGFAVSACHV 422
QY 361 H 361
DB 423 H 423

RESULT 12
AAU06617
ID AAU06617 standard: Protein: 459 AA.

XX AAU06617;

DT 24-OCT-2001 (first entry)

XX Human-pro-Asp 2(a) delta TM (His)6.

XX Human: Aspartyl protease; beta-secretase; nootropic; ASP2;
KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
KW amyloid-beta; Abeta; Human-pro-Asp 2(a) delta TM (His)6; mutant; mutain.
XX
OS Homo sapiens.
OS Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..21

FT Protein /label= Signal_peptide

FT MISC-difference /label= Mature_Human_pro-Asp_2(a)_delta_TM_(His)6

FT MISC-difference 214 /note= "Encoded by CAC"

FT MISC-difference 454..459 /note= "Encoded by CAGCAGCAGCAGCAGCAG"

FT Region 454..459

FT /label= His-tag

XX /note= "Nickel binding region to aid purification"

XX W0200149098-A2.

PD 12-JUL-2001.

XX 09-MAY-2001; 2001WO-1B00798.

PR 09-MAY-2001; 2001WO-1B00798.

XX (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

XX (YANR/) YAN R.

XX Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;

PI WPI: 2001-502549/55.

DR N-PSDB; AAS11531.

XX Novel purified polypeptide comprising fragment of mammalian aspartyl

PT protease 2, lacking Asp2 transmembrane domain and retaining beta

PT secretase activity of Asp2 useful for identifying inhibitors of Asp2

PT activity

PS Claim 149; Fig 12; 185pp; English.

XX The invention relates to a purified polypeptide comprising a fragment of

CC mammalian aspartyl protease (Asp2) protein which lacks the Asp2

CC transmembrane domain and the Asp2 protein, and where the polypeptide and

CC the fragment retain the beta-secretase activity of the mammalian Asp2

CC protein. The invention also details polynucleotides for the Asp

CC proteins and vectors expressing them, and a polypeptide (isoform of

CC amyloid protein precursor (APP) comprising the amino acid sequence of an

CC APP or its fragment containing an APP cleavage site recognizable by a

CC mammalian beta-secretase, and further comprising two lysine residues at

CC the carboxyl terminus of the amino acid sequence of the mammalian APP or

CC APP fragment. Also included in the invention are methods of identifying

CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are

CC useful for treating Alzheimer's disease. APP is useful in methods for

CC identifying inhibitors or modulators of human Asp2 activity and

CC amyloid-beta (Abeta) peptide production. APP is also useful in designing

CC therapeutics for the treatment or prevention of Alzheimer's disease.

CC APP comprising the APP-Sw-beta-secretase peptide sequence (NLDN), which

CC is associated with increased levels of Abeta processing is useful in

CC assays relating the Alzheimer's research. The expression vector is useful

CC for recombinantly expressing APP. Nucleic acids that hybridize to

CC Asp oligonucleotides are useful as probes or primers. The probes are

CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in

CC Northern and Southern blots. The present sequence is Human-pro-

CC Asp 2(a) delta TM (His)6 protein, which lacks the C-terminal

CC transmembrane domain and has a His tag to aid purification.

SO Sequence 459 AA:

Query Match 100.0%; Score 1907; DB 22; Length 459;
Best Local Similarity 100.0%; Pred. No. 1,4e-191;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDNLRGKSGGGYVEMTGSPPTNTLIVDTGSSNFVGAAPHPLHRYRQRLSTYR 60
DB 63 MYDNLRGKSGGGYVEMTGSPPTNTLIVDTGSSNFVGAAPHPLHRYRQRLSTYR 122
QY 61 DLKRGVYVYTGKMGELGTDLVSIPIGPNVTYRANIAATTESDKFINSNMEGITGL 120
DB 123 DLKRGVYVYTGKMGELGTDLVSIPIGPNVTYRANIAATTESDKFINSNMEGITGL 182
QY 121 AYAETARPDDSLPEPFDSLVKQTHVFNLFSLQCGAGFPLNQSEVLASVGSMTIGIDH 180
DB 183 AYAETARPDDSLPEPFDSLVKQTHVFNLFSLQCGAGFPLNQSEVLASVGSMTIGIDH 242
QY 181 SLYTGSIMWTPIRREMYEVIIVREINQDLMCKEYNYDKSIYDSGTTLRLPKKVF 240
DB 243 SLYTGSIMWTPIRREMYEVIIVREINQDLMCKEYNYDKSIYDSGTTLRLPKKVF 302
QY 241 EAAVSKIAASSTEFKPDGFWLGEOLVQWAGTTPMNIIPVLSILMGVTVNQSFRITL 300

```

Db      303 EAAKSIKAASTKFPDGFVLGEOLVCWQAGTTPWNIPEYISLYLMGEVTNQSFRITLL 362
QY      301 POQYLRPVEDVATSDDDCYKFAISOSSTGTVMGAVIMEGFYVFPDRARRIGFAVSACHV 360
Db      363 POQYLRPVEDVATSDDDCYKFAISOSSTGTVMGAVIMEGFYVFPDRARRIGFAVSACHV 422
QY      361 H 361
Db      423 H 423

RESULT 13
AA007216
ID      AA007216 standard; Protein; 459 AA.
XX
AC      AA007216;
XX
DT      24-OCT-2001 (first entry)
XX
DE      Human aspartyl protease 2a deltatm (His)6.
XX
KW      Human: aspartyl protease 1; Asp-1; neurotrophic; neuroprotective;
KW      aspartyl protease 2; Asp2; amyloid protein precursor; App;
KW      beta-secretase; Alzheimer's disease; HuAsp-2adeltatm (His)6.
XX
OS      Homo sapiens.
XX
FH      Key
FH      Peptide
FT      1..21
FT      /note= "Signal peptide"
FT      22..459
FT      /note= "Mature human aspartyl protease 2a deltatm (His)6"
FT      Misc-difference 454
FT      /note= "Encoded by cag"
FT      Misc-difference 455
FT      /note= "Encoded by cag"
FT      Misc-difference 456
FT      /note= "Encoded by cag"
FT      Misc-difference 457
FT      /note= "Encoded by cag"
FT      Misc-difference 458
FT      /note= "Encoded by cag"
FT      Misc-difference 459
FT      /note= "Encoded by cag"
FT      /note= "Encoded by cag"
XX
PN      WO200149097-A2.
XX
PD      12-JUL-2001.
XX
PF      09-MAY-2001; 2001WO-1B00797.
XX
PR      09-MAY-2001; 2001WO-1B00797.
XX
PA      (BIEN/) BIENKOWSKI M J.
PA      (GURNEY/) GURNEY M E.
PA      (HEIN/) HEINRIKSON R L.
PA      (PARO/) PARODI L A.
PA      (YANR/) YAN R.
XX
PI      Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;
XX
XX      WPI: 2001-502548/55.
XX      N-PSDB: AAS11716.
XX
PT      Novel purified polypeptide comprising fragment of mammalian aspartyl
PT      protease 2, lacking Asp2 transmembrane domain and retaining beta
PT      secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT      activity
XX
PS      Claim 149; Fig 12; 185pp; English.
XX
CC      The invention relates to a novel purified polypeptide comprising a

```

```

CC      fragment of mammalian aspartyl protease 2 (Asp2) protein which lacks the
CC      Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide
CC      and the fragment retain the beta-secretase activity of the mammalian Asp2
CC      protein. Also included is an isoform of amyloid protein precursor (APP)
CC      comprising the amino acid sequence of a APP or its fragment containing
CC      an APP cleavage site recognisable by a mammalian beta-secretase, and
CC      further comprising two lysine residues at the carboxyl terminus of the
CC      amino acid sequence of the mammalian APP or APP fragment. The
CC      polypeptides are used for assaying for modulators of beta-secretase
CC      activity; identifying agents that inhibit the APP processing activity
CC      of human Asp2 aspartyl protease (Hu-Asp2); identifying agents that
CC      modulate the activity of Asp2; and for reducing cellular production of
CC      amyloid beta (Abeta) from APP. Agents identified by the above methods
CC      are useful for treating Alzheimer's disease; and for identifying
CC      modulators of amyloid-beta (Abeta) peptide production, for use in
CC      designing therapeutics for the treatment or prevention of Alzheimer's
CC      disease. Probes and primers derived from Asp nucleic acid sequences
CC      are useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC      Northern and Southern blots. The present sequence represents the amino
CC      acid sequence of human Asp-2a delta TM (His)6 construct which has
CC      a 6 histidine tag and lacks the transmembrane domain. This construct was
CC      used for expression and purification of human Asp2a in insect cells.
XX
SQ      Sequence 459 AA:

```

```

Query Match          100.0%; Score 1907; DB 22; Length 459;
Best Local Similarity 100.0%; Pred. No. 1,4e-191;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MYDNLRGKSGGQYVEMVGSPPQTLNLIVDTGSSNFAVGAAPHFLLRRYOROLSTSYR 60
Db      63 MYDNLRGKSGGQYVEMVGSPPQTLNLIVDTGSSNFAVGAAPHFLLRRYOROLSTSYR 122
QY      61 DLRKGVYVPTGKMGELGDTLVSLPHGPNTVYANLAIIESDKFLINGSNMGIIIGL 120
Db      123 DLRKGVYVPTGKMGELGDTLVSLPHGPNTVYANLAIIESDKFLINGSNMGIIIGL 182
QY      121 AYAETARPDDSLPEFPDLSLVKQTHVNLFSIQLCGAGPPLNQEVLASVSGSMIIIGIDH 180
Db      183 AYAETARPDDSLPEFPDLSLVKQTHVNLFSIQLCGAGPPLNQEVLASVSGSMIIIGIDH 242
QY      181 SLYTGSLWTPTRRREYVIVIVREINQDLMKCKEYNYKSLVDSGTTNLRPKYF 240
Db      243 SLYTGSLWTPTRRREYVIVIVREINQDLMKCKEYNYKSLVDSGTTNLRPKYF 302
QY      241 EAAVSIKAASTKFPDGFVLGEOLVCWQAGTTPWNIPEYISLYLMGEVTNQSFRITLL 300
Db      303 EAAVSIKAASTKFPDGFVLGEOLVCWQAGTTPWNIPEYISLYLMGEVTNQSFRITLL 362
QY      301 POQYLRPVEDVATSDDDCYKFAISOSSTGTVMGAVIMEGFYVFPDRARRIGFAVSACHV 360
Db      363 POQYLRPVEDVATSDDDCYKFAISOSSTGTVMGAVIMEGFYVFPDRARRIGFAVSACHV 422
QY      361 H 361
Db      423 H 423

RESULT 14
AAE02595
ID      AAE02595 standard; Protein; 459 AA.
XX
AC      AAE02595;
XX
DT      10-AUG-2001 (first entry)
XX
DE      Human-Asp-2(a) deltatm (His)6 protein.
XX
KW      Human: alpha-secretase; amyloid precursor protein; App; therapy;
KW      Alzheimer's disease; antialzheimer's; aspartyl protease 2a; Asp2a;
KW      beta-secretase; Asp-2a delta TM; histidine tag; mutant; mutein.
XX
OS      Homo sapiens.

```

```

OS Synthetic.
XX Key Location/Qualifiers
FH MISC-difference 214
FT MISC-difference 454 /note= "Encoded by CAG"
FT MISC-difference 454 /note= "Encoded by CAG"
FT MISC-difference 455 /note= "Encoded by CAG"
FT MISC-difference 456 /note= "Encoded by CAG"
FT MISC-difference 456 /note= "Encoded by CAG"
FT MISC-difference 457 /note= "Encoded by CAG"
FT MISC-difference 458 /note= "Encoded by CAG"
FT MISC-difference 459 /note= "Encoded by CAG"
FT MISC-difference 459 /note= "Encoded by CAG"
XX WO200123533-A2.
XX 05-APR-2001.
XX 22-SEP-2000; 2000WO-US26080.
XX 23-SEP-1999; 99US-0155493.
XX 23-SEP-1999; 99WO-US20881.
XX 13-OCT-1999; 99US-0416901.
XX 06-DEC-1999; 99US-0169232.
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX Gurney M, Bienkowski MJ;
XX WPI: 2001-290516/30.
XX N-PSDB: AAD06753.
XX Enzymes that cleave the alpha-secretase site of the amyloid precursor
XX protein, useful for the treatment of Alzheimer's disease -
XX Example 10; Fig 12; 189pp; English.
XX The present invention relates to enzymes for cleaving the alpha-
XX secretase site of the amyloid precursor protein (APP) and methods of
XX identifying those enzymes. The methods may be used to identify enzymes
XX that may be used to cleave the alpha-secretase cleavage site of the APP
XX protein. The enzymes may be used to treat or modulate the progress of
XX Alzheimer's disease. The present sequence is human Aspartyl protease 2a
XX (Asp 2a) deltatm (His)6 protein which is obtained by deleting the C-
XX transmembrane domain and adding a histidine tag at the C-terminal end.
XX This sequence has beta-secretase protease activity.
XX
XX Sequence 459 AA:
XX
XX Query Match 100.0%; Score 1907; DB 22; Length 459;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-191;
XX Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MVDNLRKSGSGGYVEMTVGSPPTLNLIVDTGSSNFAVGAAPHPFLHRYQROLSTYR 60
XX 63 MVDNLRKSGSGGYVEMTVGSPPTLNLIVDTGSSNFAVGAAPHPFLHRYQROLSTYR 122
XX
XX 61 DLRKGVYPTTQKGWEGELGTDLVSIPIHGPNTVYRANIAATTESKFFINSGNMEGIIGL 120
XX
XX 123 DLRKGVYPTTQKGWEGELGTDLVSIPIHGPNTVYRANIAATTESKFFINSGNMEGIIGL 182
XX
XX 121 AYAELIARPDSDLPEPFDLVKQTHVNLFSIQLGAGFPLNQSEVLASVSGSMITIGIDH 180
XX 183 AYAELIARPDSDLPEPFDLVKQTHVNLFSIQLGAGFPLNQSEVLASVSGSMITIGIDH 242
XX
XX 181 SLVYGSLLWYTPIRRMWYEVIIVREINIGDOLKMDCKEYNNDKSIYVSGTINLRPKKVF 240
XX 243 SLVYGSLLWYTPIRRMWYEVIIVREINIGDOLKMDCKEYNNDKSIYVSGTINLRPKKVF 302

```

```

XX 241 EAAVSKRAASSTKEKFPDGFMLGEOLVCMQAGTTPMNIIPVLSLIMGEVYNQSPRITL 300
XX
XX 303 EAAVSKRAASSTKEKFPDGFMLGEOLVCMQAGTTPMNIIPVLSLIMGEVYNQSPRITL 362
XX
XX 301 PQQYLRPEVDVATSDDCYKFAISQSTGTVMGAVIMESGFYVFPDRAKRIRIGFAVSACHV 360
XX 363 PQQYLRPEVDVATSDDCYKFAISQSTGTVMGAVIMESGFYVFPDRAKRIRIGFAVSACHV 422
XX
XX 361 H 361
XX 423 H 423
XX
XX RESULT 15
XX AAB07898
XX ID AAB07898 standard; Protein; 460 AA.
XX
XX AC AAB07898;
XX
XX 14-NOV-2000 (first entry)
XX
XX Amino acid sequence of a human beta-secretase enzyme fragment.
XX
XX Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
XX amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
XX inhibitor.
XX
XX Homo sapiens.
XX
XX WO200047618-A2.
XX 17-AUG-2000.
XX
XX 10-FEB-2000; 2000WO-US03819.
XX
XX 10-FEB-1999; 99US-0119571.
XX 15-JUN-1999; 99US-0139172.
XX
XX (ELAN-) ELAN PHARM INC.
XX
XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
XX Slnha S, Tatsuno G, Tung J, Wang S, McConlogue LJ;
XX WPI: 2000-533011/48.
XX
XX Purified beta-secretase protein used in assays to discover inhibitors
XX which can be used for the treatment of amyloidogenic diseases e.g.
XX Alzheimer's disease -
XX
XX Claim 55; Fig 3A; 121pp; English.
XX
XX The specification describes a beta-secretase enzyme. The enzyme cleaves
XX beta-amyloid precursor protein to produce beta-amyloid peptide. This
XX enzyme is therefore implicated in the production of amyloid plaque
XX components which accumulate in the brains of individuals afflicted with
XX Alzheimer's disease. Inhibitors of beta-secretase are administered to
XX a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
XX disease-like pathology to test if they maintain or improve cognitive
XX ability or reduce the plaque burden. The compounds are used for the
XX treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
XX present sequence represents a human beta-secretase enzyme fragment.
XX
XX Sequence 460 AA:
XX
XX Query Match 100.0%; Score 1907; DB 21; Length 460;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-191;
XX Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MVDNLRKSGSGGYVEMTVGSPPTLNLIVDTGSSNFAVGAAPHPFLHRYQROLSTYR 60
XX 63 MVDNLRKSGSGGYVEMTVGSPPTLNLIVDTGSSNFAVGAAPHPFLHRYQROLSTYR 122
XX
XX 61 DLRKGVYPTTQKGWEGELGTDLVSIPIHGPNTVYRANIAATTESKFFINSGNMEGIIGL 120

```

```

Db      |||||||
123 DLKGVVPTQGWEGELGTDLVSIPHGPNVTYRANIAAITESDKFFINGSMWEGILGL 182
QY      121 AYAETARPDDSLPEPFDLSLVKOTHPNLFSLQCGAGFPLNOSSEVLASVGSMIIGIDH 180
      |||||||
Db      183 AYAETARPDDSLPEPFDLSLVKOTHPNLFSLQCGAGFPLNOSSEVLASVGSMIIGIDH 242
QY      181 SLYTGSLWYTPIRREMYEVIIVRVEINGODLKMDCKEYNTDKSIVDSGTNLRLPKKVF 240
      |||||||
Db      243 SLYTGSLWYTPIRREMYEVIIVRVEINGODLKMDCKEYNTDKSIVDSGTNLRLPKKVF 302
QY      241 EAAVKSIRAASTSEKFPDGFMLGEOLVCMQAGTTPWNIEPVISLYLMGEVTNOSFRITIL 300
      |||||||
Db      303 EAAVKSIRAASTSEKFPDGFMLGEOLVCMQAGTTPWNIEPVISLYLMGEVTNOSFRITIL 362
QY      301 POQYLRPEVDVATSDODCYKFAISOSSTGTVMGAVIMEGFYVFDRAKRRIQFAVSACHV 360
      |||||||
Db      363 POQYLRPEVDVATSDODCYKFAISOSSTGTVMGAVIMEGFYVFDRAKRRIQFAVSACHV 422
QY      361 H 361
      |
Db      423 H 423

```

Search completed: October 30, 2002, 12:27:05
 Job time : 145.577 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:23:30 ; Search time 69.1843 Seconds
(without alignments)
501.389 Million cell updates/sec

Title: US-09-724-571-75

Perfect score: 1907
Sequence: 1 MVDNLRRKSGGGYVEMTVG.....VVEDRRKRIGFAVSACHVH 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.71.*

1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1907	100.0	501	2 A59090	aspartic proteinase
2	308	16.2	384	2 JC7574	pepsinogen A - Afr
3	308	16.2	385	2 JC7575	pepsinogen A - bul
4	306	16.0	387	2 B38302	pepsin (EC 3.4.23.
5	303	15.9	383	2 JC7573	pepsinogen C - Afr
6	302	15.8	388	1 S19682	pepsin A (EC 3.4.2
7	301	15.8	382	1 PECH	pepsin A (EC 3.4.2
8	298.5	15.7	383	2 A14443	pepsin (EC 3.4.23.
9	298.5	15.7	396	2 A34401	cathepsin E (EC 3.
10	296	15.5	387	2 C38302	pepsin (EC 3.4.23.
11	296	15.5	391	2 A43356	cathepsin E (EC 3.
12	295.5	15.5	384	2 A39314	pepsin A (EC 3.4.
13	295.5	15.5	412	1 KHRUD	cathepsin D (EC 3.
14	295	15.5	387	2 D38302	pepsin (EC 3.4.23.
15	287	15.0	407	1 KHRUD	cathepsin D (EC 3.
16	286	15.0	387	2 E38302	pepsin (EC 3.4.23.
17	285.5	15.0	398	2 S66465	cathepsin E (EC 3.
18	285.5	15.0	444	2 T24204	cathepsin E (EC 3.
19	285	14.9	398	1 S19684	hypothetical prote
20	285	14.9	398	1 S19684	pepsin A (EC 3.4.2
21	283	14.8	388	2 E51185	cathepsin D (EC 3.
22	281.5	14.8	381	1 CSMHB	pepsin A (EC 3.4.2
23	281	14.7	388	1 PEMOAJ	chymosin (EC 3.4.2
24	279.5	14.7	410	1 KHRUD	pepsin A (EC 3.4.2
25	278.5	14.6	386	1 PECH	pepsin A (EC 3.4.2
26	278	14.6	387	2 JC7245	pepsinogen A - com
27	278	14.6	388	1 PEHU	pepsin A (EC 3.4.2
28	277.5	14.6	380	2 E47176	chymosin (EC 3.4.2
29	277.5	14.6	396	2 S36865	cathepsin E (EC 3.

30	277	14.5	388	2 A30142	pepsin A (EC 3.4.2
31	275	14.4	388	2 B30142	pepsin A (EC 3.4.2
32	273.5	14.3	381	1 CMO	chymosin (EC 3.4.2
33	272	14.3	389	2 JE0371	pepsin C (EC 3.4.2
34	270.5	14.2	377	1 PEMOCT	pepsinogen C (EC 3.4
35	270.5	14.2	389	2 A38302	pepsin (EC 3.4.23.
36	270	14.2	376	2 I45856	aspartic proteinase
37	268.5	14.1	344	1 KHRGD	cathepsin D (EC 3.
38	267.5	14.0	381	2 JC7247	prochymosin - comm
39	266	13.9	405	2 A25379	saccharopepsin (EC
40	265.5	13.9	380	2 S03433	cathepsin D (EC 3.
41	264	13.8	396	2 J47207	aspartic proteinase
42	263.5	13.8	388	2 JC7246	pepsinogen C - com
43	262.5	13.8	394	2 B43356	pepsinogen C (EC 3.4
44	261.5	13.7	387	2 A45117	cathepsin (EC 3.4
45	261.5	13.7	388	2 A29937	aspartic proteinase

ALIGNMENTS

RESULT 1

A59090 aspartic proteinase (EC 3.4.23.-) BACE precursor - human

N:Alternate names: beta-secretase; beta-site APP cleaving enzyme

C:Species: Homo sapiens (man)

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-May-2000

C:Accession: A59090

R:Vaasat, R.; Bennett, B.D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E.A.; Denis, P.; Teplio

M.A.; Biere, A.L.; Curran, E.; Burgess, T.; Louis, J.C.; Collins, F.; Treanor, J.; Ro

Science 286, 735-741, 1999

A:Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the tran

A:Reference number: A59090; MIM:20002972

A:Note: submitted to GenBank, September 1999

A:Accession: A59090

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-501 <VAS>

A:Cross-references: GB:AF190725; NID:96118538; PIDN:AF04142.1; PID:96118539

C:Genetics:

A:Gene: BACE

C:Superfamily: beta-secretase

C:Keywords: Alzheimer's disease; aspartic proteinase; brain; glycoprotein; hydrolase;

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-45/Domain: propeptide #status predicted <PRO>

F:46-501/Product: acid proteinase BACE #status predicted <MAT>

F:461-477/Domain: transmembrane #status predicted <TRN>

F:99,289/Active site: Asp #status predicted

F:153,172,223,354/Binding site: carbohydrate (asn) (covalent) #status predicted

F:330-380/Disulfide bonds: #status predicted

Query Match 100.0%; Score 1907; DB 2; Length 501;

Best Local Similarity 100.0%; Pred. No. 5, 5e-154;

Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDNLRRKSGGGYVEMTVGSPPTQTLNIVDTGSSNFAVGAAPFLHRYQRLSTYR 60

DB 63 MVDNLRRKSGGGYVEMTVGSPPTQTLNIVDTGSSNFAVGAAPFLHRYQRLSTYR 122

QY 61 DLKRGVVPYVTVGGWEGELGTLVSIHPGPNVYRANIAATTESDKFFINGSNWEGITGL 130

DB 123 DLKRGVVPYVTVGGWEGELGTLVSIHPGPNVYRANIAATTESDKFFINGSNWEGITGL 182

QY 121 AYAEIARPDLSLEPFPSLVKQTHVPLVFSLOLCGAEPLNOSVVLASVGSMTIGIDH 180

DB 183 AYAEIARPDLSLEPFPSLVKQTHVPLVFSLOLCGAEPLNOSVVLASVGSMTIGIDH 242

QY 181 SLVYGLWYPIPIREMYEYIVLVEINGDLKMDCKEYNDKSIYDSGTTNRLPKVF 240

DB 243 SLVYGLWYPIPIREMYEYIVLVEINGDLKMDCKEYNDKSIYDSGTTNRLPKVF 302

QY 241 EAAVSKIAASSTREKPFDFGWLGEOLVCMQAGTTPNPIFVYISLYLMGEVNTNOSFRITIL 300

Db 303 EAAVKSIRKAASSTKPPDGFGLGEOLVCMQAGTTPNNIFPVISLVMGEVNTQSFRTIL 362
QY 301 PQOYLPRVEDVATSDCCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKIGAVSACHV 360
|||
Db 363 PQOYLPRVEDVATSDCCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKIGAVSACHV 422
QY 361 H 361
|
Db 423 H 423

RESULT 2

JC7574
pepsinogen A - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C:Accession: JC7574; PC7119
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens
A:Reference number: JC7573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: JC7574
A:Molecule type: mRNA
A:Residues: 1-384 <IKU>
A:Cross-References: DDBJ:AB045380
A:Accession: PC7119
A:Molecule type: protein
A:Residues: 16-35;57-76 <IK2>
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like
C:Genetics:
A:Gene: Pga
A:Superfamily: pepsin
C:Keywords: stomach; zymogen

Query Match 16.2%; Score 308; DB 2; Length 384;
Best Local Similarity 27.5%; Pred. No. 2.6e-18;
Matches 99; Conservative 60; Mismatches 133; Indels 68; Gaps 16;

QY 13 YVEMTVGSPPTQTLNIVDTGSSNFVAGAPHPFL-----HRYRQLSSTYRDLRKG 65
|||
Db 72 YGTTISIGTPQSTFYVIEDTGSANLWV---PSVYSSQACSHNNFNPOQSTFGATNTP 128
QY 66 VVPTQGWEGELGTDVLSIPHGPVTVRANIAITESDK-FFINGSMWEGILGLAYAE 124
|||
Db 129 VSIQGTGSMGFLGVDYQV---GNIQSNQMGFLSESEPSFLYSPFDILGLAFPS 185
QY 125 IARPDDSLPEFDSLVKQTHVP-NLFSIQLCGAGFLPNSGVLASVGSMTIGIDHSLY 183
|||
Db 186 IA--SSQATPVFDMMWSQGLIPQNLFSVYLSDDG-----QTGSYVLFGVDNSLY 233
QY 184 TGSIMWTPIRREMYEVIIIVRVEINGDGL-KMDCKEYNDKSIYDSGTNLRLPKVPE 241
|||
Db 234 SGLNMPVLAETLYWITLIDSVISINQVACSSC-----QALVDTGSLMTGPETPT- 286
QY 242 AAVKSIKAASTKPPDGFGLGEOLV-CMQAGTTPNNIFPVISLVMGEVNTQSFRTIL 300
|||
Db 287 ANIQNTIGASQDSN-----GQYVINCNNISNMPITV-----TIN 321
QY 301 PQOYLPRVEDVATSDCCYK-FAISQSSTGT-----VMGAVIMEGYVVFDRARRKIGFA 354
|||
Db 322 GVOYPLSPSAYVNOGCGSSGFQAMNLPNLSGDLMLIGDVFIRQYFYFDRANNYVAIA 381

RESULT 3

JC7575
pepsinogen A - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C:Accession: JC7575
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens

A:Reference number: JC7573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: JC7575
A:Molecule type: mRNA
A:Residues: 1-385 <IKU>
A:Cross-References: DDBJ:AB045376
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like
C:Genetics:
A:Gene: Pga
A:Superfamily: pepsin
C:Keywords: stomach; zymogen

Query Match 16.2%; Score 308; DB 2; Length 385;
Best Local Similarity 27.8%; Pred. No. 2.6e-18;
Matches 99; Conservative 64; Mismatches 133; Indels 60; Gaps 15;

QY 13 YVEMTVGSPPTQTLNIVDTGSSNFVAG---AAPHPFLRRYQRLSSTYRDLRKGYYV 68
|||
Db 73 YGTTISIGTPQSTFYVIEDTGSANLWVPSVYSSQACSHNNFNPOQSTFGATNTPVSI 132
QY 69 PYTQGWEGELGTDVLSIPHGPVTVRANIAITESDK-FFINGSMWEGILGLAYAEIAR 127
|||
Db 133 QYGTGSMGFLGVDYQV---GNIQNTQIFGLSQSPGSGFLYSPFDILGLAFPSLA- 188
QY 128 PDDSLPEFDSLVKQTHVP-NLFSIQLCGAGFLPNSGVLASVGSMTIGIDHSLYTGS 186
|||
Db 189 -SSQATPVFDMMWSQGLIPQDLFSVYLSQSG-----QS-----GSFVLFQGVDTSTYTG 237
QY 187 LMYPIRREMYEVIIIVRVEINGDGLK--DCKEYNDKSIYDSGTNLRLPKVFEAAV 244
|||
Db 238 LNMVPLRAETLYWITLIDSVISINQVACSSC-----SALVDTGSLMTGPETPT- 287
QY 245 KSIKAASTKPPDGFGLGEOLVCMQAGTTPNNIFPVISLVMGEVNTQSFRTILPQOY 304
|||
Db 288 ANIQYIGANDQSNQYV---INCNNISNMPITV-----TINQY 326
QY 305 LRPVED-VATSDDC---YKFAISQSSTGT---VMGAVIMEGYVVFDRARRKIGFA 354
|||
Db 327 PLPASAYVROSQOSTSGFQAMNLPSTSGDLMLIGDVFIRRYVYFDRANNYVMA 382

RESULT 4

B38302
pepsin (EC 3.4.23.-) II-1 precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Jun-1991 #sequence_revision 20-Sep-1991 #text_change 23-Feb-1997
C:Accession: B38302
R:Kageyama, T.; Tanabe, K.; Koizumi, O.
J. Biol. Chem. 265, 17031-17038, 1990
A:Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nu
A:Reference number: A38302; MUID:91009127
A:Accession: B38302
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-387 <KAG>
A:Cross-References: GB:M59235; GB:J05638
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion

Query Match 16.0%; Score 306; DB 2; Length 387;
Best Local Similarity 27.1%; Pred. No. 3.9e-18;
Matches 98; Conservative 67; Mismatches 131; Indels 66; Gaps 15;

QY 13 YVEMTVGSPPTQTLNIVDTGSSNFVAG---AAPHPFLRRYQRLSSTYRDLRKGYYV 68
|||
Db 75 YGTTISIGTPQSTFYVIEDTGSANLWVPSVYSSQACSHNNFNPOQSTFGATSETLSI 134
QY 69 PYTQGWEGELGTDVLSIPHGPVTVRANIAITESD---KFFINGSMWEGILGLAYAEI 125
|||
Db 135 TYTGSGMTGILGIDYKV---GNIEPTNQIFGLSKTEPITFLV--APPDGILGLAYPSI 189
QY 126 ARPDDSLPEFDSLVKQTHV-NLFSIQLCGAGFLPNSGVLASVGSMTIGIDHSLYT 184
|||

```
Db      365 EYSSVYDLGNQVGFTAA 383
```

```

Db      133 VSIYTGSGSMGILGYLTGVK---GGISDTNQIFELSETPEGEFLYFAPFDGILLAAAPS 189
QY      125 IARPDSDLPEPDSLVNVOHY-PNLFSLQLCGAGFPILNQSEVLASVGGSMITGIDHSLY 183
Db      190 IS--SSCATPEFDINMORLVSQDLFSYLTISAD---DQS-----GSVIFGGIDSSY 237
QY      184 TGLMTYPIREWEYEVIIYVEIINGDL--KMDKEYNDKSIYDSGTJNLRLKKYFE 241
Db      238 TGSLSMWVPVSEGYWQSLVSDISITMNGKTIACAKG-----QAYDITSLTGTSPFA 291
QY      242 AAKSKTAASTKEKFPDGFMLGEOLV-CWQAGTPPMNIFPYSIYLMEVTNQSPRITIL 300
Db      292 NIOSDGASNSD-----GEKVSCSAISLPDIYF-----TIN 325
QY      301 POOY-LKRPVEDVATSDODCK-----FAISGSSGTGVAGVIMEGFYVFDRAKRISFA 354
Db      326 GVOYPLPFPSPAYIILOSQSGCTSGFGKMDVPTESEGLMIIGDVFIRQYFVFDRAANQVGLA 385

RESULT 7
PECH
pepsin A (EC 3.4.23.1) precursor - chicken
N:Alternate names: pepsinogen A
C:Species: Gallus gallus (chicken)
C:Date: 18-Apr-1984 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C:Accession: J030370. A00984
R:Sakamoto, N.; Saiga, H.; Yasugi, S.
Biochem. Biophys. Res. Commun. 250, 420-424, 1998

```


Db 384 VGLA 387

RESULT 12

A39314

gastricsin (EC 3.4.23.3) precursor - bullfrog

C:Species: Rana catesbeiana (bullfrog)

C:Date: 19-Jun-1992 #sequence,revision 19-Jun-1992 #text,change 22-Jun-1999

C:Accession: A39314

R:Yakabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Miki, K.; Kurokawa, K.; Ito, H.; Kageya

J. Biol. Chem. 266, 22436-22443, 1991

A:Title: Purification, characterization, and amino acid sequences of pepsinogens and pep

A:Reference number: A39314; MUID:92042186

A:Accession: A39314

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-384 <YAK>

A:Cross-references: GB:M73750; NID:g213687; PIDN:AAA49530.1; PID:g213688

A:Superfamily: pepsin

C:Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match 15.5%; Score 295.5; DB 2; Length 384;

Best Local Similarity 26.1%; Pred. No. 3e-17;

Matches 98; Conservative 59; Mismatches 130; Indels 89; Gaps 15;

QY 13 YVEMTVGSPQTLNIVDTGSSNFAVGAAPHPFL-----HRYQROLSTYRDLKRG 65

Db 67 YVGEISIGTPPQNFVLPDTGSSNLMV--PSTYCOSQACTNHPQNPQSSSSYSNNOQ 123

QY 66 VVPTGCKWEGELGMDLVSIPIHGNVYRANIA-----AITESDEFFINGSNMEGI 117

Db 124 FSLQIGTSGTILIGDYQI-----QNTAISQDEFGLSVTEPGNFEY-AQFDGI 173

QY 118 LGLAAEAIARPDSDLEPPDLSLVKQTHVPN-LFSLQCGAGFPLNQSEVLASVGSMTIG 176

Db 174 LGLAVPSIA-EGGATVWQGMIOQLINQPLFAFLSGQNSQN-----QGEVAG 223

QY 177 GIDHSIYTSMTWTPRRRWYEVIIIVREINGOD---LKMCKEKNMKSIVDSGTTNL 233

Db 224 GVDQNYISQIWTVPSTSTWQIGQGFVNGQATGWCSCG-----QGLVDGTSL 277

QY 234 RPKKFEAAVSIKAASTEFKPDGFMLEQDV-CWQAGTTPWNIIFYI-----SL 284

Db 278 TAPQSFVSLMSIQAGQDN-----GQYAVSCSNISQSLPTISFTISGVFPLPSA 329

QY 285 YLMGEVTNO---SFRITLIPQOQLRPVEVVAISQDDCYFAISQSTGVMNAVIMEGY 341

Db 330 YVLOQNSGCTIGIMPTVLPSONGQPL-----WILDVFLRQY 368

QY 342 VVFDRAKRRIKGFVSA 357

Db 369 SVYDLGNQVGFRAAA 384

RESULT 13

KHHUD

cathepsin D (EC 3.4.23.5) precursor [validated] - human

N:Alternate names: precathepsin D

C:Species: Homo sapiens (man)

C:Date: 28-Dec-1987 #sequence,revision 28-Dec-1987 #text,change 15-Sep-2000

C:Accession: A25771; S30749; PC2066; I59236; I57716

R:Fauts, P.L.; Kornfeld, S.; Chirgwin, J.M.

Proc. Natl. Acad. Sci. U.S.A. 82, 4910-4914, 1985

A:Title: Cloning and sequence analysis of cDNA for human cathepsin D.

A:Reference number: A25771; MUID:85270436

A:Accession: A25771

A:Molecule type: mRNA

A:Residues: 1-412 <FAU>

A:Cross-references: EMBL:M1233; NID:g181179; PIDN:AAB59529.1; PID:g181180

R:Westley, B.R.; May, F.E.B.

Nucleic Acids Res. 15, 3773-3786, 1987

A:Title: Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive human breasts

A:Reference number: S30749; MUID:87231068

A:Accession: S30749

A:Molecule type: mRNA

A:Residues: 1-412 <MES>

A:Cross-references: EMBL:X05344; NID:g29677; PIDN:CAA28955.1; PID:g29678

R:May, F.E.B.; Smith, D.J.; Westley, B.R.

Gene 134, 277-282, 1993

A:Title: The human cathepsin D-encoding gene is transcribed from an estrogen-regulate

A:Reference number: PC2066; MUID:94085791

A:Accession: PC2066

A:Molecule type: DNA

A:Residues: 1-23 <MA>

A:Cross-references: GB:L12980; NID:g291930; PIDN:AAA16314.1; PID:g455429

A:Experimental source: MCF-7 cell

R:Cavallies, V.; Augereau, P.; Rochefort, H.

Proc. Natl. Acad. Sci. U.S.A. 90, 203-207, 1993

A:Title: Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate

A:Reference number: I59236; MUID:93126342

A:Accession: I59236

A:Status: translation not shown; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-22 <CAV1>

A:Cross-references: GB:S5257; NID:g263124; PIDN:AAD13868.1; PID:g4261568

R:Augereau, P.; Miralles, F.; Cavallies, V.; Gaudelot, C.; Parker, M.; Rochefort, H.

Mol. Endocrinol. 8, 693-703, 1994

A:Title: Characterization of the proximal estrogen-responsive element of human cathep

A:Reference number: I57716; MUID:95021301

A:Accession: I57716

A:Status: translation not shown; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-22 <CAV2>

A:Cross-references: GB:S74689; NID:g786350; PIDN:AAD14156.1; PID:g4261856

R:Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.

submitted to the Brookhaven Protein Data Bank, April 1993

A:Reference number: A51839; PDB:1LVA

A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 65-161,170-241

A:Reference number: A51840; PDB:1LVB

A:Contents: annotation; X-ray crystallography, 2.5 angstroms, with inhibitor residues

R:Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Hosur, M.V.; Sowder II, R.C.; Cachau, R.E.;

Proc. Natl. Acad. Sci. U.S.A. 90, 6796-6800, 1993

A:Title: Crystal structures of native and inhibited forms of human cathepsin D: Impli

A:Reference number: A48229; MUID:93342076

A:Contents: annotation; X-ray crystallography, 2.5 angstroms

C:Comment: Cathepsin D is a ubiquitous lysosomal proteinase.

C:Comment: In addition to the propeptide, residues 163-168 and 411-412 are proteolyti

C:Comment: The carbohydrate bound to 134-Asn contains a mannose-6-phosphate that is b

C:Genetics:

A:Gene: GDB:CTSD

A:Cross-references: GDB:120512; OMIM:116840

A:Map position: 11p15.5-11p15.5

C:Function:

A:Description: limited specificity endopeptidase

A:Pathway: intracellular protein degradation

C:Superfamily: pepsin

C:Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradati

F:1-20/Domin: signal sequence #status predicted <SIG>

F:21-64/Domin: propeptide #status predicted <PRO>

F:65-162,169-410/Product: cathepsin D #status experimental <MAT>

F:267,329-356/Region: phosphotransferase recognition

F:91-160,110-117,286-290,329-366/Disulfide bonds: #status experimental

F:97,295/Active site: Asp #status experimental

F:134,263/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 15.5%; Score 295.5; DB 1; Length 412;

Best Local Similarity 28.5%; Pred. No. 3e-17;

Matches 105; Conservative 58; Mismatches 146; Indels 59; Gaps 15;

QY 13 YVEMTVGSPQTLNIVDTGSSNFAVGAAPHPFL-----HRYQROLSTYRDLKRG 66

Db 79 YVGEISIGTPPQNFVLPDTGSSNLMVPSIHCKLIDIAICWIIHKYNSDKSSTYVKNGTSF 138

RESULT 14
D38302
pepsin (EC 3.4.23.-) II-4 precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Jun-1991 #sequence_revision 20-Sep-1991 #text_change 23-Feb-1997
R:Accession: D38302
R:Kageyama, T.; Tanabe, K.; Koiwai, O.
J. Biol. Chem. 265, 17031-17038, 1990
A:Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nucle
A:Reference number: A38302; MUID:91009127
A:Accession: D38302
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-387 <KAG>
A:Cross-references: GB:M59235; GB:J05638
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion

QY 236 PKVFEAAVKSIIKAASSTEKEFPDGFWLGEQLV-CWQAGTTPWNIFPVISLYLMGEVTNQ 294

```

Db      299  PVDEKELQKAIGAVPLIQ-----GEYMIPEKVS-----LPITFKLGQ----- 340
QY      295  FRITILPOOYLRPVEDVATSODDCYKFAIS-----QSTGTVMGAVIMEGFYVFD 345
Db      341  -NYELHPEKYLKVSQAGKT-----ICLSGFMGMDIPPSGPLWILGDVFIQCYTYVFD 393
QY      346  RAKKRIGPVA 357
Db      394  REYNREVFKA 405

```

Search completed: October 30, 2002, 12:31:28
 Job time : 70.1843 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:20:50 ; Search time 3.20393 Seconds
(without alignments)
277.344 Million cell updates/sec

Title: US-09-724-571-81

Perfect score: 34

Sequence: 1 EVMXAEF 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	94.1	8	21	AA07872
2	27	79.4	7	21	AA07871
3	27	79.4	347	22	AA03166
4	26	76.5	107	20	AA135564
5	26	76.5	216	21	AA195715
6	26	76.5	218	22	AA03744
7	26	76.5	254	20	AA152214
8	26	76.5	271	20	AA15227
9	26	76.5	271	21	AA08205
10	26	76.5	271	22	AA064357
11	26	76.5	271	22	AA092851

12	26	76.5	271	22	AA088446
13	26	76.5	276	21	AA042611
14	26	76.5	282	21	AA057116
15	26	76.5	559	17	AA090619
16	26	76.5	579	22	AA096355
17	26	76.5	587	21	AA096355
18	26	76.5	587	21	AA096355
19	26	76.5	587	21	AA096355
20	26	76.5	821	20	AA096355
21	26	76.5	869	20	AA096355
22	26	76.5	1249	22	AA071313
23	25	73.5	14	21	AA07888
24	25	73.5	27	22	AA028205
25	25	73.5	27	22	AA03380
26	25	73.5	27	22	AA018839
27	25	73.5	27	22	AA05165
28	25	73.5	27	22	AA06559
29	25	73.5	27	22	AA014432
30	25	73.5	27	22	AA026845
31	25	73.5	27	22	AA002159
32	25	73.5	54	19	AA079151
33	25	73.5	54	20	AA081408
34	25	73.5	88	22	AA056968
35	25	73.5	135	22	AA026145
36	25	73.5	144	22	AA031774
37	25	73.5	152	22	AA096397
38	25	73.5	189	22	AA052817
39	25	73.5	225	22	AA065926
40	25	73.5	237	22	AA090704
41	25	73.5	280	22	AA090751
42	25	73.5	443	21	AA042806
43	25	73.5	453	21	AA068799
44	25	73.5	479	20	AA093603
45	25	73.5	482	22	AA031776

ALIGNMENTS

RESULT 1

ID AAB07872 standard; peptide; 8 AA.

AC AAB07872;

XX 14-NOV-2000 (first entry)

DE A beta-secretase inhibitor peptide.

KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
amyloid plaque component; Alzheimer's disease; amyloidogenic disease;

KW inhibitor.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 4 /note="hydroxyethylene"

XX WO200047618-A2.

XX 17-AUG-2000.

XX 10-FEB-2000; 2000WO-US03819.

XX 10-FEB-1999; 99US-0119571.

XX 15-JUN-1999; 99US-0139172.

XX (ELAN-) ELAN PHARM INC.

PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Shiba S, Tatsuno G, Tung J, Wang S, McConlogue L;

Human membrane or
Human ORFX ORF2375
Human prostate can
Sulfolobus solfata
Putative P. abyss
Neisseria meningit
Neisseria meningit
Porphyromonas ging
Porphyromonas ging
A peptide fragment
Human peptide #856
Peptide #882 encod
Protein #838 encod
Human brain expres
Human bone marrow
Peptide #866 encod
Peptide #882 encod
Peptide #841 encod
Receptor protein t
Receptor protein t
Protonbacterium
Novel human diagno
Novel human secret
Escherichia coli p
Drosophila melanog
C glutamicum prote
Human ORFX ORF2570
Human acid sequenc
Tomato Xa21 clone
Novel human secret

DR WPI: 2000-533011/48.
XX
XX Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
PS Disclosure; Page 12; 121pp; English.
XX
CC The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents an inhibitor of beta-secretase enzyme.
XX
SQ Sequence 8 AA:

Query Match 94.1%; Score 32; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 6,4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVMXVAEF 8
1 |||||||
Db 1 EVMXVAEF 8

RESULT 2
AAB07871
ID AAB07871 standard; peptide; 7 AA.
AC
XX AAB07871;
XX
DT 14-NOV-2000 (first entry)
XX
DE A beta-secretase inhibitor peptide.
XX
XX Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 3 /note="hydroxyethylene"
FT
XX
XX WO200047618-A2.
XX
PD 17-AUG-2000.
XX
XX 10-FEB-2000; 2000WO-US03819.
PF
XX 10-FEB-1999; 99US-0119571.
PR
XX 15-JUN-1999; 99US-0139172.
PR
XX (ELAN-) ELAN PHARM INC.
PA
XX
XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX
XX WPI: 2000-533011/48.
XX
XX Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
XX Disclosure; Page 12; 121pp; English.
XX
XX The specification describes a beta-secretase enzyme. The enzyme cleaves

CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents an inhibitor of beta-secretase enzyme.
XX
SQ Sequence 7 AA:

Query Match 79.4%; Score 27; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 6,4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VMXVAEF 8
1 |||||||
Db 1 VMXVAEF 7

RESULT 3
AAG93168
ID AAG93168 standard; Protein; 347 AA.
AC
XX AAG93168;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 6922.
XX
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN Ep1108790-A2.
XX
PD 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-0127688.
PF
XX 16-DEC-1999; 99JP-0377484.
PR
XX 07-APR-2000; 2000JP-0159162.
PR
XX 03-AUG-2000; 2000JP-0280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI: 2001-376931/40.
XX
DR N-PSDB; AAH68387.
XX
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
XX Claim 17; SEQ ID NO: 6922; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Corynebacterium bacterium, and identifying a homologue of a gene derived
CC from Corynebacterium bacterium. Corynebacterium bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the

CC European Patent Office.

SQ Sequence 347 AA:

Query Match 79.4%; Score 27; DB 22; Length 347;

Best Local Similarity 62.5%; Pred. No. 1.8e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVMXVAEF 8
1:1111
78 EILIVAEF 85RESULT 4
AAV35564

ID AAV35564 standard; Protein; 107 AA.

XX AAV35564;

AC AAV35564;

DT 13-SEP-1999 (first entry)

DE Chlamydia pneumoniae protein not found in C. trachomatis.

KM Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;

KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;

XX vaccine; neutralising epitope.

OS Chlamydia pneumoniae.

PM WO9927105-A2.

XX 03-JUN-1999.

PF 20-NOV-1998; 98WO-IB01890.

XX 04-NOV-1998; 98US-0107078.

PR 21-NOV-1997; 97ER-0014673.

XX (GEST) GENSET.

PA Grifffals R;

PI WPI: 1999-357842/30.

DR Genome sequence of Chlamydia pneumoniae

PT Page 1307; Disclosure; 1912pp; English.

XX AAY34584-Y35879 represent the proteins encoded by all the open reading

CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.

CC C. pneumoniae causes respiratory disease such as pneumonia and

CC bronchitis and is thought to be a contributing factor in heart

CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema

CC nodosum or pharyngitis. The polypeptides encoded by the open reading

CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in

CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae

CC nucleotide sequences can also be used as immunogenic compositions,

CC especially where the vector directs the expression of a neutralising

CC epitope of C. pneumoniae.

CC Sequence 107 AA;

SQ Query Match 76.5%; Score 26; DB 20; Length 107;

Best Local Similarity 62.5%; Pred. No. 93;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVMXVAEF 8
111:111
91 EVMRIARF 98

Db 91 EVMRIARF 98

RESULT 5
AAY95715

ID AAY95715 standard; Protein; 216 AA.

XX AAY95715;

AC 25-OCT-2000 (first entry)

DT Cosmid cHRIM5 encoded protein P21-7r.

DE Cosmid cHRIM5; nematocide; nematode; biological control agent;

KM transgenic plant; helminthiasis; P21-7r.

XX Xenorhabdus bovienii.

OS WO200042855-A1.

PM 27-JUL-2000.

XX 24-JAN-2000; 2000WO-GB00219.

XX 22-JAN-1999; 99GB-0001499.

PR (HORT-) HORTICULTURE RES INT.

XX Morgan JAW, Jarrett P, Ellis D, Ousley MA;

DR WPI: 2000-499157/44.

XX N-PSDB; AAA50029.

PT Novel composition used to control parasitic nematodes, especially in

XX plants such as maize, cotton, soya, and rice, comprises a bacterium

XX which is a symbiont of an entomopathogenic nematode -

XX Example 6; Page 44; 74pp; English.

XX The present sequence is that of protein P21-7r encoded by an open

XX reading frame identified in cosmid cHRIM5 (see AAA50029). cHRIM5 was

XX obtained by ligating Xenorhabdus bovienii strain I73 (NCIMB 40986)

XX CC Sau3A-digested DNA fragments into the BamHI site of the Stratagene

XX CC cosmid vector SupercosI, packaging into Escherichia coli XL Blue I,

XX CC and screening for nematocidal activity against Caenorhabditis elegans.

XX CC Analysis of the DNA indicated a number of open reading frames for

XX CC which the corresponding protein sequences were determined (see

XX CC AAY95685-Y95735). Nematodes can be controlled through the use of

XX CC bacteria associated symbiotically with an entomopathogenic nematode.

XX CC Such bacteria include Xenorhabdus and Photobacterium spp. such as X.

XX CC bovienii strain I73. The symbiotic bacteria, an engineered

XX CC bacterium, or a nematocidal protein obtained from such bacteria,

XX CC can be used to control helminthiasis in a human or domesticated

XX CC animal or for the control of plant pathogen nematodes. Also

XX CC claimed are vectors for expressing nematocidal proteins in host

XX CC cells, and transgenic plants.

XX Sequence 216 AA;

SQ Query Match 76.5%; Score 26; DB 21; Length 216;

Best Local Similarity 62.5%; Pred. No. 1.9e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVMXVAEF 8
1:1111
108 EIMGVANF 115

Db 108 EIMGVANF 115

RESULT 6
AAB63744

ID AAB63744 standard; Protein; 218 AA.

XX AAB63744;

AC 26-MAR-2001 (first entry)

DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1106.

KW Human: breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WC200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US14749.
 XX
 PR 28-MAY-1999; 99US-0136526.
 PR 10-SEP-1999; 99US-0153454.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Obata Y;
 XX
 DR WPI: 2001-025274/03.
 XX
 PT Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 PS Example 1; Page 696-697; 799pp; English.
 XX
 CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP), respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer.
 XX
 SQ Sequence 218 AA;
 XX
 Query Match 76.5%; Score 26; DB 22; Length 218;
 Best Local Similarity 75.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 EYMXVAEF 8
 |||
 Db 167 EVKNVAEF 174

XX
 PA (GEST) GENSET.
 XX
 PI Griffais R;
 XX
 DR WPI: 1999-357842/30.
 XX
 PT Genome sequence of Chlamydia pneumoniae
 XX
 PS Page 1062-1063; Disclosure; 1912pp; English.
 XX
 CC AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAY31990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading
 CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotide sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.
 XX
 SQ Sequence 254 AA;
 XX
 Query Match 76.5%; Score 26; DB 20; Length 254;
 Best Local Similarity 50.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 1 EYMXVAEF 8
 |||
 Db 204 ELLATAEF 211

RESULT 8
 AAY15227
 ID AAY15227 standard; protein: 271 AA.
 XX
 AC AAY15227;
 XX
 DT 26-OCT-1999 (first entry)
 XX
 DE Human receptor protein (HURP) 6 amino acid sequence.
 XX
 KW receptor; cancer; autoimmune disorder; inflammation;
 KW antagonist; cell surface protein; cell signalling;
 KW antibody; human receptor protein; HURP; reproductive disorder;
 KW developmental disorder; gastrointestinal disorder.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Misc-difference 63
 FT /note= "Potential CAMP-/cGMP-dependent protein-
 FT kinase phosphorylation site"
 FT
 FT Misc-difference 95
 FT /note= "Potential casein kinase II-
 FT phosphorylation site"
 FT
 FT Misc-difference 114
 FT /note= "Potential casein kinase II-
 FT phosphorylation site"
 FT
 FT Misc-difference 213
 FT /note= "Potential casein kinase II-
 FT phosphorylation site"
 FT
 FT Misc-difference 6
 FT /note= "Potential protein kinase C-
 FT phosphorylation site"
 FT
 FT Misc-difference 25
 FT /note= "Potential protein kinase C-
 FT phosphorylation site"
 FT
 FT Misc-difference 59
 FT /note= "Potential protein kinase C-
 FT phosphorylation site"

FT	Misc-difference	63	/note=	"Potential protein kinase C-phosphorylation site"
FT				
FT	Misc-difference	75	/note=	"Potential protein kinase C-phosphorylation site"
FT				
FT	Misc-difference	123	/note=	"Potential protein kinase C-phosphorylation site"
FT				
FT	Misc-difference	135	/note=	"Potential protein kinase C-phosphorylation site"
FT				
FT	Misc-difference	189	/note=	"Potential protein kinase C-phosphorylation site"
FT				
FT	Misc-difference	203	/note=	"Potential protein kinase C-phosphorylation site"
FT				
FT	Misc-difference	247	/note=	"Potential protein kinase C-phosphorylation site"
FT				
FT	Misc-difference	160	/note=	"Potential tyrosine kinase-phosphorylation site"
FT				
FT	Binding-site	71..78	/note=	"Potential ATP/GTP-binding site-motif A (P-loop)"
FT				
PN	WO9941375-A2.			
XX				
PD	19-AUG-1999.			
XX				
PF	05-FEB-1999;	99WO-US02572.		
XX				
PR	12-FEB-1998;	98US-0022939.		
XX				
PA	(INCYTE) INCYTE PHARM INC.			
XX				
PI	Au-Young JI, Bandman O, Baughn M, Corley NC, Guegler KJ,			
PI	Hillman JL, Lal P, Shah P, Tang YT, Yue H;			
XX				
DR	WPI: 1999-494536/41.			
DR	N-PsDB; AA06371.			
XX				
PT	New human receptor proteins, used e.g. to treat, prevent and			
PT	diagnose gastrointestinal and developmental disorders - and related			
PT	nucleic acids, vectors, transformed cells, antibodies, agonists and			
PT	antagonists			
XX				
PS	Claim 1; Page 83; 94pp; English.			
XX				
CC	The Human receptor protein 6 (HURP-6) and mouse signal recognition			
CC	particle beta subunit share 90% identity.			
CC	HURP-6 is expressed in cancerous, inflamed, reproductive and gastro-			
CC	intestinal tissue. HURP-6 therefore appears to have a role in			
CC	cancer, autoimmune/inflammatory disorders, reproductive disorders,			
CC	and gastrointestinal disorders.			
CC	This gives rise to the possibility of using an antagonist or an antibody			
CC	of HURP to treat or prevent cancer or autoimmune/inflammatory disorders.			
XX				
SQ	Sequence	271 AA;		

Query Match	76.5%	Score 26;	DB 20;	Length 271;
Best Local Similarity	75.0%	Pred. No. 2.4e+02;		
Matches	6;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;
QY	1	EYMXVAEF	8	
Db	151	EYKDYAEF	158	
RESULT 9				
AAB28205				

ID	AA28205 standard; Protein; 271 AA.
XX	
AC	AA28205;
XX	
DT	30-JAN-2001 (first entry)
XX	
DE	Novel human protein #3.
XX	
KW	Cytostatic; vaccine; human; breast tumour; antigen; breast cancer.
XX	
OS	Homo sapiens.
XX	
PN	W0200052165-A2.
XX	
PD	08-SEP-2000.
XX	
PF	29-FEB-2000; 2000MO-US05431.
XX	
PR	04-MAR-1999; 99US-0262505.
PR	19-MAR-1999; 99US-0272886.
PR	17-SEP-1999; 99US-0396313.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Lodes MJ;
XX	
DR	WPI; 2000-572184/53.
DR	N-PSDB; AAC69684.
XX	
PT	Breast tumor antigen polypeptides and polynucleotides, useful for
PT	manufacturing vaccines and compositions for treating, diagnosing, and
PT	monitoring breast cancer -
XX	
PS	Example; Fig 2; 140bp; English.
XX	
CC	The present invention relates to immunogenic portions of new human
CC	breast tumor antigens (AA28183-B28214) and their coding sequences
CC	(AAC69645-C69804). The breast tumor antigen polypeptides of the present
CC	invention and their coding sequences are useful for inhibiting the
CC	development of breast cancer in a patient. The breast tumor antigen
CC	polypeptides and polynucleotides may be used in vaccines and
CC	pharmaceutical compositions for treating breast cancer, and for
CC	diagnosing and monitoring the cancer. The present sequence is a
CC	immunogenic portion for one such human breast cancer tumor antigen.
XX	
Sequence	271 AA;

Query Match	76.5%;	Score 26;	DB 21;	Length 271;
Best Local Similarity	75.0%;	Pred. No. 2.4e+02;		
Matches	6;	Conservative	0;	Mismatches 2;
			Indels	0;
Qy	1 EVMXVAEF 8			
Db	151 EVKDYAEF 158			
RESULT 10				
AAG64357				
ID	AAG64357	standard; Protein; 271 AA.		
XX				
AC	AAG64357;			
XX				
DT	01-OCT-2001	(first entry)		
XX				
DE		Human signal recognition particle receptorbeta.		
XX				
KW		Human; signal recognition particle receptor beta; SRPBeta.		
XX				
OS		Homo sapiens.		
XX				
PN	CNL279290-A.			
XX				
PD	10-JAN-2001.			

XX 23-JUN-1999; 99CN-0108547.
 XX 23-JUN-1999; 99CN-0108547.
 XX (UYFU-) UNIV FUDAN.
 XX Yu L, Fu Q, Zhao Y;
 DR WPI: 2001-266742/28.
 DR N-PSDB; AAH73878.
 XX New human signal recognition particle receptor beta nucleic acid for
 PT preparing the protein encoded by it -
 XX Claim 2; Pages 15-16 (Disclosure): 20pp; Chinese.
 XX The present sequence is the protein sequence for human signal recognition
 CC particle receptorbeta (SRPBeta). The present protein is the homolog of
 CC mouse SRPBeta. Application of human SRPBeta coding sequence and
 CC protein, and their preparing process are also disclosed.

XX Sequence 271 AA;
 SQ
 Query Match 76.5%; Score 26; DB 22; Length 271;
 Best Local Similarity 75.0%; Pred. NO. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 EVMXVAAF 8
 || ||||
 DB 151 EYKDYAEF 158

RESULT 11
 AAB92851
 ID AAB92851 standard; Protein: 271 AA.
 XX AAB92851;
 AC 26-JUN-2001 (first entry)
 DT Human protein sequence SEQ ID NO:11414.
 XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 OS EP1074617-A2.
 XX EP1074617-A2.
 PN 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-0116126.
 PF 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-318749/34.
 DR Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length CDNs defined in the specification, and for the detection
 PT full-length CDNs -
 PS Claim 8; SEQ ID 11414; 2537pp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602
 CC full-length CDNs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to a
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are also useful for the full-length
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length CDNs. The primers allow obtaining of the full-length
 CC CDNs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human amino acid sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 271 AA;
 SQ
 Query Match 76.5%; Score 26; DB 22; Length 271;
 Best Local Similarity 75.0%; Pred. NO. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 EVMXVAAF 8
 || ||||
 DB 151 EYKDYAEF 158

RESULT 12
 AAB88446
 ID AAB88446 standard; Protein: 271 AA.
 XX AAB88446;
 AC 23-MAY-2001 (first entry)
 DT Human membrane or secretory protein clone PSEC0230.
 XX Human; secretory protein; membrane protein; vaccine; gene therapy;
 KW Rheumatoid arthritis; diabetes.
 OS Homo sapiens.
 XX EP1067182-A2.
 PN 10-JAN-2001.
 PD 07-JUL-2000; 2000EP-0114090.
 PF 08-JUL-1999; 99JP-0194179.
 XX 11-JAN-2000; 2000JP-0118775.
 PR 02-MAY-2000; 2000JP-0183766.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
 PI WPI: 2001-093989/11.
 DR N-PSDB; AAF93873.
 XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -
 PS Claim 1; SEQ ID 260; 609pp + CD ROM; English.
 CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by

CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA
 CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretion
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to
 CC identify modulators (agonists and antagonists) of expression and
 CC activity. The antibodies and antagonists may also be used as therapeutic
 CC agents to down regulate expression and activity. The antibodies may also
 CC be used as diagnostic agents for detecting the presence of the
 CC polypeptides in samples (e.g. by enzyme linked immunosorbant assay
 CC (ELISA). Examples of diseases which may be treated include rheumatoid
 CC arthritis and diabetes.

XX Sequence 271 AA;

Query Match 76.5%; Score 26; DB 22; Length 271;
 Best Local Similarity 75.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVKXVAEF 8
 II IIII
 DB 151 EVKDAEF 158

RESULT 13

AAB42611
 ID AAB42611 standard; Protein; 276 AA.

XX AAB42611;

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF2375 polypeptide sequence SEQ ID NO:4750.

Human: Open reading frame; ORFX; detection; cytosolic; hepatotropic;
 human; vulnery; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
 anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
 immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 hypotensive; dermatologic; immunosuppressive; antineumatic; antihypertensive;
 antiviral; antibacterial; antifungal; antipneumatic; antihypertensive;
 antineumatic; gene therapy; cancer; proliferative disorder; hypertension;
 neurodegenerative disorder; osteoarthritis; graft vs host disease;
 cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 cholesterol ester storage; systemic lupus erythematosus; infection;
 severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 bone damage; cartilage damage; antinflammatory disease; coagulation;
 thrombosis; contraceptive.

XX Homo sapiens.

OS WO200058473-A2.

XX 05-OCT-2000.

PD 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0340763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;
 XX WPI; 2000-602362/57.
 DR N-PSDB; AAC76820.

PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

PS Claim 11; Page 3927; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosolic; hepatotropic; vulnery;
 CC antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressive;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatologic; antineumatic; antipneumatic;
 CC antihypertensive; antibacterial; antiviral; antifungal; antineumatic;
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 276 AA;

Query Match 76.5%; Score 26; DB 21; Length 276;
 Best Local Similarity 75.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVKXVAEF 8
 II IIII
 DB 156 EVKDAEF 163

RESULT 14

AAB57116
 ID AAB57116 standard; Protein; 282 AA.

XX AAB57116;

DT 13-MAR-2001 (first entry)

XX Human prostate cancer antigen protein sequence SEQ ID NO:1694.

Human: prostate cancer; prostate cancer antigen; detection; diagnosis;
 human; vulnery; antiproliferative; cytosolic; hepatotropic; neuroprotective;
 anticonvulsant; osteopathic; antiparkinsonian; neurotrophic; neuroprotective;
 antidiabetic; hypotensive; dermatologic; antineumatic; antipneumatic;
 antihypertensive; antibacterial; antiviral; antifungal; antineumatic;
 wound; infectious disease.

XX Homo sapiens.

OS WO200055174-A1.

XX 21-SEP-2000.

PD 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

PR (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.
 XX Rosen CA, Ruben SM;
 XX WPI: 2000-587513/55.
 DR N-PSDB; AAF16319.
 XX Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -
 XX
 XX Claim 11: Page 2165-2166; 2338pp; English.
 PS
 XX AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardiactive, immunomodulatory, muscular, vulnery, gastrointestinal,
 CC nephrotropic, antineoplastic, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.
 CC
 XX Sequence 282 AA:
 SQ
 XX
 XX Query Match 76.5%; Score 26; DB 21; Length 282;
 Best Local Similarity 75.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EVKXVAEF 8
 I I I I I I
 Db 162 EVKXVAEF 169

RESULT 15
 AAR90619
 ID AAR90619 standard; Protein; 559 AA.
 XX
 AC AAR90619;
 XX
 DT 29-JUN-1996 (first entry)
 XX
 DE Sulfolobus solfataricus amylase for alpha, alpha-trehalose produ.
 XX
 XX transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;
 KW malto-oligosaccharide; hydrolysis.
 XX
 XX Sulfolobus solfataricus.
 OS
 XX
 PN W09534642-A.
 XX
 PD 21-DEC-1995.
 XX
 PF 14-JUN-1995; 95WO-JP01189.
 XX
 PR 21-APR-1995; 95JP-0120673.
 PR 15-JUN-1994; 94JP-0133354.
 PR 18-AUG-1994; 94JP-0194223.
 PR 31-OCT-1994; 94JP-0290394.
 PR 21-NOV-1994; 94JP-0286917.
 PR 21-NOV-1994; 94JP-0311185.
 XX
 PA (KIRI) KIRIN BEER KK.
 XX
 XX WPI: 1996-049671/05.
 DR N-PSDB; AAT12325.
 XX
 PT Sulfolobus spp. derived transferase and amylase - for production of

PT alpha, alpha-trehalose from malto-oligosaccharide(s)
 XX
 XX Claim 108, Page 235-240; 357pp; Japanese.
 PS
 XX The amylase is derived from Sulfolobus solfataricus. The amylase acts
 CC on a saccharide having at least three sugar units, which are pref.
 CC glucose units at the reducing end (the linkage between the first and
 CC second glucose units is alpha-1, alpha-1, while the linkage between the
 CC second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4
 CC linkages within the sugar chain, yielding alpha, alpha-trehalose and also
 CC mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of
 CC 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85
 CC deg.C. It has an isoelectric point of 4.3-5.4 and retains at least
 CC 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited
 CC by 5 mM copper sulphate. Use of a transferase and the amylase in
 CC succession on suitable substrates such as malto-oligosaccharides, is
 CC useful for the production of alpha, alpha-trehalose.
 CC
 XX Sequence 559 AA:
 SQ
 XX
 XX Query Match 76.5%; Score 26; DB 17; Length 559;
 Best Local Similarity 50.0%; Pred. No. 5.2e+02;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EVKXVAEF 8
 I I I I I I
 Db 138 EIMPIAQF 145

Search completed: October 30, 2002, 12:27:09
 Job time : 5.20393 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:24:20 : Search time 1.14005 Seconds
(without alignments)
171.400 Million cell updates/sec

Title: US-09-724-571-81

Perfect score: 34

Sequence: 1 EVMXVAEF 8

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	76.5	559	4	US-09-242-690A-15
2	25	73.5	54	2	US-08-456-647B-18
3	25	73.5	54	2	US-08-237-401A-18
4	25	73.5	405	4	US-09-291-023A-20
5	25	73.5	485	2	US-08-446-803-2
6	25	73.5	485	2	US-08-861-837-2
7	25	73.5	485	3	US-08-600-656-2
8	25	73.5	485	4	US-09-170-670-2
9	25	73.5	485	4	US-09-170-670-8
10	25	73.5	485	4	US-09-193-068-2
11	25	73.5	485	4	US-09-193-068-8
12	25	73.5	485	4	US-09-183-412-2
13	25	73.5	485	4	US-09-183-412-8
14	25	73.5	485	4	US-09-264-097-5
15	25	73.5	485	4	US-09-354-191A-2
16	25	73.5	513	2	US-08-459-346-19
17	25	73.5	513	2	US-07-989-847-8
18	25	73.5	513	3	US-08-889-419-19
19	25	73.5	513	3	US-08-469-411-8
20	25	73.5	513	3	US-08-889-419-19
21	25	73.5	513	6	PCT-US93-07189-19
22	25	73.5	556	3	US-08-505-377-1
23	25	73.5	556	3	US-08-798-269-1
24	25	73.5	556	4	US-09-055-210-1
25	25	73.5	816	1	US-07-640-029-1
26	25	73.5	817	1	US-07-640-029-2
27	24	70.6	57	4	US-08-630-915A-209

28	24	70.6	161	4	US-08-858-207A-284	Sequence 284, App
29	24	70.6	329	1	US-08-230-047-7	Sequence 7, Appl
30	24	70.6	341	1	US-08-314-309A-19	Sequence 19, Appl
31	24	70.6	524	3	US-08-557-210A-3	Sequence 3, Appl
32	24	70.6	539	3	US-08-557-210A-4	Sequence 4, Appl
33	24	70.6	539	3	US-08-557-210A-5	Sequence 5, Appl
34	24	70.6	691	5	PCT-US91-08442-2	Sequence 2, Appl
35	24	70.6	758	1	US-07-736-250-16	Sequence 16, Appl
36	24	70.6	972	3	US-08-335-844A-24	Sequence 24, Appl
37	23	67.6	41	1	US-08-096-741-1	Sequence 46, Appl
38	23	67.6	54	2	US-08-456-647B-46	Sequence 47, Appl
39	23	67.6	54	2	US-08-456-647B-47	Sequence 46, Appl
40	23	67.6	54	2	US-08-237-401A-46	Sequence 47, Appl
41	23	67.6	54	2	US-08-237-401A-47	Sequence 46, Appl
42	23	67.6	83	1	US-08-096-741-7	Sequence 47, Appl
43	23	67.6	215	1	US-08-431-080-22	Sequence 22, Appl
44	23	67.6	215	2	US-08-938-534-22	Sequence 22, Appl
45	23	67.6	236	4	US-09-605-858-34	Sequence 34, Appl

ALIGNMENTS

```

RESULT 1
US-09-242-690A-15
: Sequence 15, Application US/09242690A
: Patent No. 6284534
:
: GENERAL INFORMATION:
: APPLICANT: KONDO, KEIJI
: APPLICANT: MIURA, YUTAKA
: TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
: FILE OF INVENTION: SAME
: FILE REFERENCE: 049441/0118
: CURRENT APPLICATION NUMBER: US/09/242,690A
: CURRENT FILING DATE: 1999-02-23
: PRIOR APPLICATION NUMBER: PCT/JP97/02924
: PRIOR FILING DATE: 1997-08-22
: PRIOR APPLICATION NUMBER: JP 8/241062
: NUMBER OF SEQ ID NOS: 66
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 15
: LENGTH: 559
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein
US-09-242-690A-15
Query Match          76.5%; Score 26; DB 4; Length 559;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY      1 EVMXVAEF 8
Db      138 EIMPIAOF 145

RESULT 2
US-08-456-647B-18
: Sequence 18, Application US/08456647B
: Patent No. 581516
:
: GENERAL INFORMATION:
: APPLICANT: Lemke Ph.D. et al., Greg E.
: TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
: NUMBER OF SEQUENCES: 54
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: US
: ZIP: 92037

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,647B
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/237,401
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 07251/007002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-647B-18

```

Query Match	73.5%	Score 25;	DB 2;	Length 54;
Best Local Similarity	50.0%	Pred. No. 20;		
Matches	4;	Conservative	3;	Mismatches
			1;	Indels
			0;	Gaps
OY	1	EYMXVYAEF	8	
	:	:	:	
Db	8	DYMKIADF	15	

RESULT 3
 US-08-237-401A-18
 Sequence 18, Application US/08237401A
 Patent No. 5837448
 GENERAL INFORMATION:
 APPLICANT: Lemke, Ph. D. et al., Greg E.
 TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
 NUMBER OF SEQUENCES: 54
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: US
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/237,401A
 FILING DATE: 02-MAY-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/884,486
 FILING DATE: 15-MAY-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Haile Ph.D., Lisa A.
 REGISTRATION NUMBER: 38,347
 REFERENCE/DOCKET NUMBER: 07251/007001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 678-5070
 TELEFAX: (619) 678-5099

```

: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 54 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-237-401A-18

Query Match          73.5%  Score 25;  DB 2;  Length 54;
Best Local Similarity 50.0%  Pred. No. 20;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy      1 EVNXYAEF 8
       :||:|:|
Db      8 DVMKTADE 15

```

```

RESULT 4
US-09-291-023A-20
; Sequence 20: Application US/09291023A
; Patent No. 6309871
; GENERAL INFORMATION:
; APPLICANT: Outtrup, Helle
; APPLICANT: Borchert, Torben
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nielsen, Vibeke
; APPLICANT: Hoeck, Lisbeth
; TITLE OF INVENTION: Polypeptides Having Alkaline Alpha-amylase Activity And Nuclei
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 5821.010-US
; CURRENT APPLICATION NUMBER: US/09/291.023A
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: DK 1999 00438
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 405
; TYPE: prt
; ORGANISM: Bacillus
US-09-291-023A-20

```

Query Match	Score 25;	DB 4;	Length 405;
Best Local Similarity	62.5%;	Pred. No. 1.7e+02;	
Matches	5;	Conservative	1; Mismatches 2; Indels 0; Gaps 0;
QY	1 EVMKVAEF 8		
	:- :		
Db	180 EFMVAEAF 187		

RESULT 5
 US-08-446-803-2
 ; Sequence 2, Application US/08446803
 ; Patent No. 5824531
 ; GENERAL INFORMATION:
 ; APPLICANT: Otturup, Helle
 ; APPLICANT: Bisgard-Frantzen, Henrik
 ; APPLICANT: Ostergaard, Peter Rahbek
 ; APPLICANT: Rasmussen, Michael Dolberg
 ; APPLICANT: van Der Zee, Pia
 ; TITLE OF INVENTION: Alkaline Bacillus Amylase
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: No. 5824531disk of No. 58245311th American
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,803
FILING DATE: 01-June-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 4157.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-803-2

Query Match 73.5%; Score 25; DB 2; Length 485;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVXKVAEF 8
I: |||||
DB 260 EMPVAEAF 267

RESULT 6
US-08-861-837-2
Sequence 2, Application US/0861837
Patent No. 5856164
GENERAL INFORMATION:
APPLICANT: Orlup, Helle
APPLICANT: Bisgard-Frantzen, Henrik
APPLICANT: Ostergaard, Peter Rahbek
APPLICANT: Rasmussen, Michael Dolberg
APPLICANT: Van Der Zee, Pia
TITLE OF INVENTION: Alkaline Bacillus Amylase
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 58561640 No. 5856164disk of No. 5856164th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/861,837
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,803
FILING DATE: 01-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 4157.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-861-837-2

Query Match 73.5%; Score 25; DB 2; Length 485;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVXKVAEF 8
I: |||||
DB 260 EMPVAEAF 267

RESULT 7
US-08-600-656-2
Sequence 2, Application US/08600656
Patent No. 6093562
GENERAL INFORMATION:
APPLICANT: Bisgard-Frantzen, Henrik
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: AMYLASE VARIANTS
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60935620 No. 6093562disk of No. 6093562th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,656
FILING DATE: 13-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4318.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-600-656-2

Query Match 73.5%; Score 25; DB 3; Length 485;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVXKVAEF 8
I: |||||
DB 260 EMPVAEAF 267

RESULT 8
US-09-170-670-2
Sequence 2, Application US/09170670
Patent No. 6187576
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Borchert, Torben
APPLICANT: Bisgard-Frantzen, Henrik

```

; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 5276.200-US
; CURRENT APPLICATION NUMBER: US/09/170,670
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 1172/97
; EARLIER FILING DATE: 1997-10-13
; EARLIER APPLICATION NUMBER: 60/063,306
; EARLIER FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-170-670-2

Query Match
Best Local Similarity 73.5%; Score 25; DB 4; Length 485;
                          62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVMXVAEF 8
   | : ||||
Db 260 EMFAVAEF 267

RESULT 9
US-09-170-670-8
; Sequence 8, Application US/09170670
; Patent No. 6187576
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben
; APPLICANT: Bisgaard-Frantzen, Henrik
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 5276.200-US
; CURRENT APPLICATION NUMBER: US/09/170,670
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 1172/97
; EARLIER FILING DATE: 1997-10-13
; EARLIER APPLICATION NUMBER: 60/063,306
; EARLIER FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-170-670-8

Query Match
Best Local Similarity 73.5%; Score 25; DB 4; Length 485;
                          62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVMXVAEF 8
   | : ||||
Db 260 EMFAVAEF 267

RESULT 10
US-09-193-068-2
; Sequence 2, Application US/09193068
; Patent No. 6197565
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjærulff, Søren
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Andersen, Carsten
; TITLE OF INVENTION: -Amylase Variants
; FILE REFERENCE: 5709.000-US
; CURRENT APPLICATION NUMBER: US/09/193,068
; CURRENT FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 485
; TYPE: PRT

```

```

; SEQ ID NO 2
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-193-068-2

Query Match
Best Local Similarity 73.5%; Score 25; DB 4; Length 485;
                          62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVMXVAEF 8
   | : ||||
Db 260 EMFAVAEF 267

RESULT 11
US-09-193-068-8
; Sequence 8, Application US/09193068
; Patent No. 6197565
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjærulff, Søren
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Andersen, Carsten
; TITLE OF INVENTION: -Amylase Variants
; FILE REFERENCE: 5709.000-US
; CURRENT APPLICATION NUMBER: US/09/193,068
; CURRENT FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-193-068-8

Query Match
Best Local Similarity 73.5%; Score 25; DB 4; Length 485;
                          62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVMXVAEF 8
   | : ||||
Db 260 EMFAVAEF 267

RESULT 12
US-09-183-412-2
; Sequence 2, Application US/09183412
; Patent No. 6204232
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kjærulff, Søren
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09/183,412
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 60/064,662
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 60/093,234
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 1240/97
; EARLIER FILING DATE: 1997-10-30
; EARLIER APPLICATION NUMBER: PA 1998 00936
; EARLIER FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 485
; TYPE: PRT

```

```
; ORGANISM: Bacillus sp.
US-09-183-412-2
Query Match          73.5%: Score 25; DB 4; Length 485;
Best Local Similarity 62.5%: Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKXVAEF 8
   1: ||||
Db 260 EMFAVAEF 267

RESULT 13
US-09-183-412-8
; Sequence 8, Application US/09183412
; Patent No. 6204232
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nielsen, Torben L.
; APPLICANT: Kjaerulf, Soren
; TITLE OF INVENTION: Alpha-Amulase Mutants
; FILE REFERENCE: 5368-200-US
; CURRENT APPLICATION NUMBER: US/09/183,412
; EARLIER FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 60/064,662
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 60/093,234
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 1240/97
; EARLIER FILING DATE: 1997-10-30
; EARLIER APPLICATION NUMBER: PA 1998 00936
; EARLIER FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-183-412-8

Query Match          73.5%: Score 25; DB 4; Length 485;
Best Local Similarity 62.5%: Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKXVAEF 8
   1: ||||
Db 260 EMFAVAEF 267

RESULT 14
US-09-264-097-5
; Sequence 5, Application US/09264097
; Patent No. 6287826
; GENERAL INFORMATION:
; APPLICANT: No. 6287826man, Barrie Edmund
; APPLICANT: Hendriksen, Hanne Vang
; TITLE OF INVENTION: Enzymatic Preparation of Glucose Syrup
; FILE REFERENCE: 5278-200-US
; CURRENT APPLICATION NUMBER: US/09/264,097
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: PA 0321/98
; EARLIER FILING DATE: 1998-03-09
; EARLIER APPLICATION NUMBER: 60/079,209
; EARLIER FILING DATE: 1998-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 485
; TYPE: PRT
```

```
; ORGANISM: Bacillus
US-09-264-097-5
Query Match          73.5%: Score 25; DB 4; Length 485;
Best Local Similarity 62.5%: Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

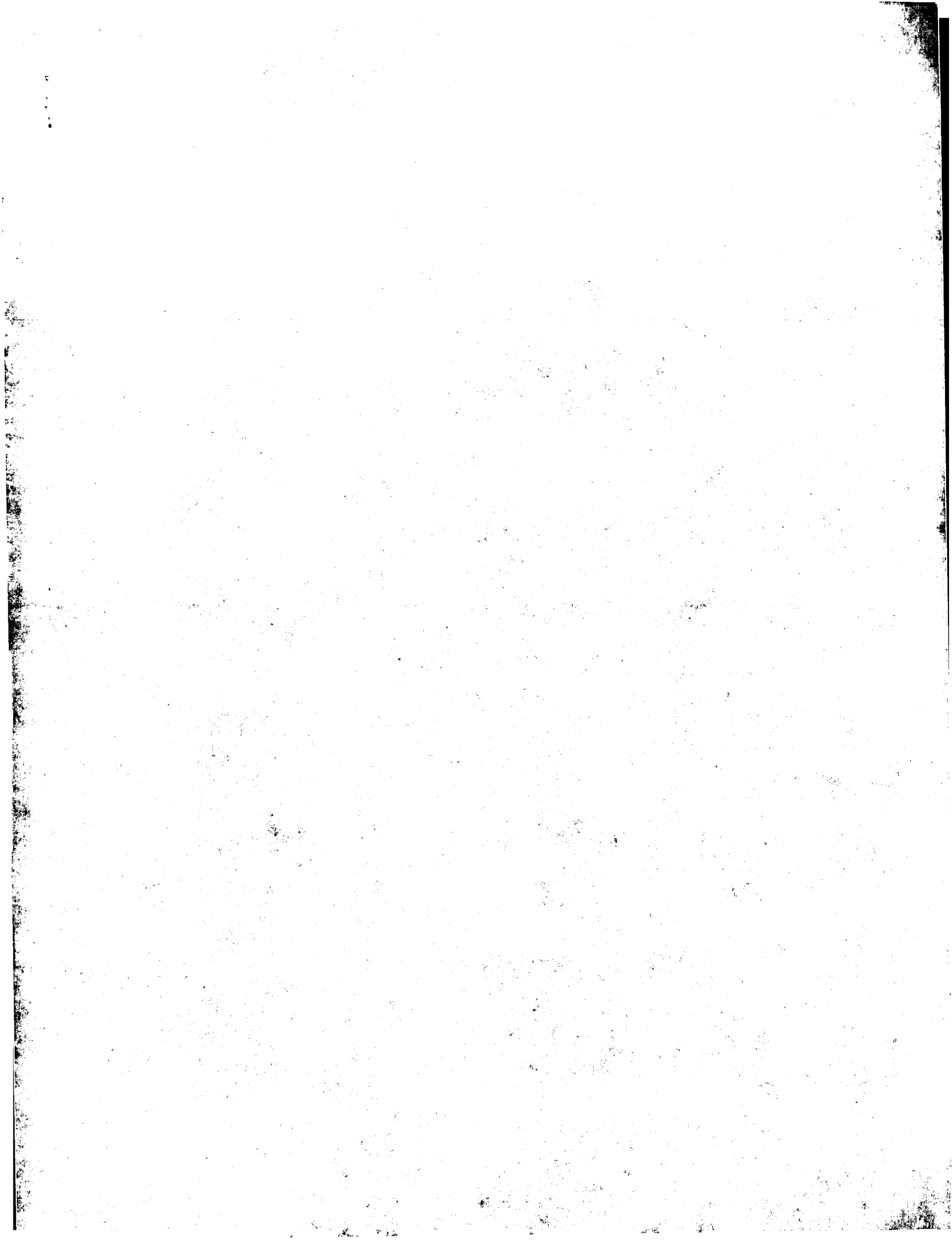
QY 1 EVKXVAEF 8
   1: ||||
Db 260 EMFAVAEF 267

RESULT 15
US-09-354-191A-2
; Sequence 2, Application US/09354191A
; Patent No. 6297038
; GENERAL INFORMATION:
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 62970380 No. 6297038disk of No. 6297038th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/354,191A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/600,656
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4318-204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-354-191A-2

Query Match          73.5%: Score 25; DB 4; Length 485;
Best Local Similarity 62.5%: Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKXVAEF 8
   1: ||||
Db 260 EMFAVAEF 267

Search completed: October 30, 2002, 12:32:34
Job time : 2.14005 secs
```



GenCore version 5.1.3
Copyright (c) 1993 - 2002, Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:23:30 ; Search time 1.5317 Seconds

(without alignments)
501.389 Million cell updates/sec

Title: US-09-724-571-81

Sequence: 1 EVWXYAEF 8

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	85.3	195	2	C97255
2	29	85.3	236	2	B70728
3	28	82.4	1256	2	A10673
4	28	82.4	1256	2	AB2042
5	28	82.4	2638	1	A42545
6	27	79.4	181	2	T11902
7	27	79.4	184	2	C68133
8	27	79.4	201	2	F69988
9	27	79.4	598	2	B71095
10	27	79.4	601	2	D96001
11	27	79.4	802	2	T05596
12	27	79.4	1560	2	T30282
13	26	76.5	46	2	B55209
14	26	76.5	150	2	B55209
15	26	76.5	223	2	B64205
16	26	76.5	224	2	S02216
17	26	76.5	252	2	E72060
18	26	76.5	252	2	B86564
19	26	76.5	269	2	A56487
20	26	76.5	278	2	C96421
21	26	76.5	324	2	T15662
22	26	76.5	324	2	B84452
23	26	76.5	490	2	E96010
24	26	76.5	498	2	H82494
25	26	76.5	533	2	A61616
26	26	76.5	533	2	A61253
27	26	76.5	543	2	G83825
28	26	76.5	558	2	J05135
29	26	76.5	561	2	S73087

30	26	76.5	562	2	G75044	acetoacetate synth
31	26	76.5	587	2	D81881	probable succinate
32	26	76.5	587	2	F81138	succinate dehydrog
33	26	76.5	603	2	D70445	aspartate--tRNA li
34	26	76.5	623	2	B82536	ABC transporter At
35	26	76.5	1014	2	C83990	beta-galactosidase
36	26	76.5	1085	2	G89056	protein K09H11.3 l
37	25	73.5	136	2	I70179	dynein-like protei
38	25	73.5	152	2	S14388	dynein-like protei
39	25	73.5	152	2	E90639	hypothetical prote
40	25	73.5	152	2	E85490	hypothetical prote
41	25	73.5	152	2	AC0067	conserved hypotnet
42	25	73.5	152	2	AD0517	conserved hypotnet
43	25	73.5	187	2	A86847	acyltransferase [l
44	25	73.5	188	2	I70180	dynein-like protei
45	25	73.5	200	2	I70169	dynein-like protei

ALIGNMENTS

RESULT 1
C97255
thymidine kinase (EC 2.7.1.21) [similarity] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2001
C:Accession: C97255
R:Noiling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A86900; MUID:21359325; PMID:21359325
A:Accession: C97255
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-195 <KUD>
A:Cross-references: GB:AE001437; PIDN:AKR0830.1; PID:G15025935; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2887
C:Superfamily: thymidine kinase
C:Keywords: phosphotransferase

Query Match 85.3%; Score 29; DB 2; Length 195;
Best Local Similarity 62.5%; Pred. No. 9;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVWXYAEF 8
DB 130 ELWMAIEF 137

RESULT 2
B70728
hypothetical protein RV2558 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70728
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Rajadna, M.A.; Rogers, K.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:9825987
A:Accession: B70728
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <COL>
A:Cross-references: GB:Z77250; GB:AL12456; NID:93261617; PIDN:CA001046.1; PID:e25533
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2558

Query Match 85.3%; Score 29; DB 2; Length 236;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVXVAF 8
 ||| |||
 DB 216 EVLDVAF 223

RESULT 3

A10673
 probable hydrolase STY1503 [imported] - Salmonella enterica subsp. enterica serovar Typh

C:Species: Salmonella enterica subsp. enterica serovar Typh

A:Note: this species has also been called Salmonella typh

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: A10673

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero

A:Reference number: AB0502; PMID:11677608

A:Accession: A10673

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-594 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD01762.1; PID:G16502610; GSPDB:GN00176

C:Gene: STY1503

C:Superfamily: trehalose trehalohydrolase

Query Match 82.4%; Score 28; DB 2; Length 594;
 Best Local Similarity 75.0%; Pred. No. 52;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVXVAF 8
 ||| |||
 DB 148 EVMPVAF 155

RESULT 4

AB2042
 hypothetical protein all1888 [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AB2042

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iritaghi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; M0ID:21595285; PMID:1175840

A:Accession: AB2042

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1256 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA073587.1; PID:G17130978; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all1888

Query Match 82.4%; Score 28; DB 2; Length 1256;
 Best Local Similarity 75.0%; Pred. No. 11e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVXVAF 8
 ||| |||
 DB 74 EVMAFAF 81

RESULT 5

A42545
 genome polypeptide - Langat virus (strain TP21) (fragment)
 N:Contains: nonstructural protein NS1; nonstructural protein NS2a; nonstructural prot

NS5

C:Species: Langat virus

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 19-Jan-2001

C:Accession: A42545; A61409; C61409

R:Jacono-Connors, L.C.; Schmaljohn, C.S.

Virology 188, 875-880, 1992

A:Title: Cloning and sequence analysis of the genes encoding the nonstructural protei

A:Reference number: A42545; M0ID:92263794

A:Accession: A42545

A:Molecule type: genomic RNA

A:Residues: 1-2638 <IAC>

A:Cross-references: GB:S53565; NID:q249315; PIDN:AA022165.1; PID:q249316

R:Gulikhoov, F.; Heinz, F.X.; Mandl, C.W.; Holzmann, H.; Kunz, C.; Gressikova, M.

J. Gen. Virol. 72, 333-338, 1991

A:Title: The relationship between the flavivirus skalia and langat as revealed by

A:Reference number: A61409; M0ID:91132129

A:Accession: A61409

A:Status: not compared with conceptual translation

A:Molecule type: genomic RNA

A:Residues: 319-337 <GU1>

A:Accession: C61409

A:Status: not compared with conceptual translation

A:Molecule type: genomic RNA

A:Residues: 877-994 <GU2>

C:Superfamily: yellow fever virus genome polypeptide

C:Keywords: glycoprotein; nonstructural protein; nucleotide binding; P-loop; polypept

F:1-352/Product: nonstructural protein NS1 #status predicted <NS1>

F:1-352/Product: nonstructural protein NS2a #status predicted <NS2a>

F:353-382/Product: nonstructural protein NS2b #status predicted <NS2b>

F:383-713/Product: nonstructural protein NS3 #status predicted <NS3>

F:714-1334/Product: nonstructural protein NS4 #status predicted <NS4>

F:912-919/Region: nucleotide-binding motif A (P-loop)

F:1335-1483/Product: nonstructural protein NS4a #status predicted <NS4a>

F:1484-1735/Product: nonstructural protein NS4b #status predicted <NS4b>

F:1736-2638/Product: nonstructural protein NS5 #status predicted <NS5>

F:85,207,223,1212,1671,1950/Binding site: carbohydrate (Asn) (covalent) #status p

Query Match 82.4%; Score 28; DB 1; Length 2638;
 Best Local Similarity 85.7%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 VNXVAF 8
 ||| |||
 DB 153 VMTVAF 159

RESULT 6

T11902
 NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain N002 - Thermus aquaticus thermophi

C:Species: Thermus aquaticus thermophilus

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: T11902

R:Yano, T.; Chu, S.S.; Sled, V.S.; Ohnishi, T.; Yagi, T.

submitted to the EMBL Data Library, March 1996

A:Description: Cloning, sequencing, and expression studies of the proton-translocatin

A:Reference number: Z17372

A:Accession: T11902

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-181 <YAN>

A:Cross-references: EMBL:U52917; NID:q1279860; PID:G1279865; PIDN:AAA97942.1

C:Genetics:

A:Gene: N002

Query Match 79.4%; Score 27; DB 2; Length 181;
 Best Local Similarity 75.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVXVAF 8
 ||| |||

Db 59 EVWGVASF 66

RESULT 7

C69133

DNA-dependent RNA polymerase, subunit E' - Methanobacterium thermoautotrophicum (strain C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999
C:Accession: C69133

R:Smith, D.R.; Doucette-Stamm, L.A.; Delouche, C.; Lee, H.; Dubois, J.; Alredge, T.;
Olin, D.; Spadafora, R.; Vicalini, R.; Mang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514
A:Accession: C69133

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-184 <MTH>

A:Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PIDN:AAB84770.1; PID:g262131
A:Experimental source: strain Delta H
C:Genetics:

A:Gene: MTH264
A:Start codon: TTG
C:Superfamily: DNA-directed RNA polymerase subunit E

Query Match
Best Local Similarity 79.4%; Score 27; DB 2; Length 184;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVWVAAF 8
11: |||
Db 88 EVIRIAEF 95

RESULT 8

F69988

hypothetical protein ytbQ - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: F69988

R:Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.: Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capano, V.; Carter, N.M.; Ch
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koettler, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

Y, M.; Ogawa, K.; Ogawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, E.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033
A:Accession: F69988

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-201 <KUN>

A:Cross-references: GB:Z99119; GB:AL009126; NID:g2635411; PIDN:CAB14996.1; PID:el185891;
A:Experimental source: strain 168
C:Genetics:

A:Gene: ytbQ

Query Match
Best Local Similarity 79.4%; Score 27; DB 2; Length 201;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVWVAAF 8
11: |||
Db 22 DIMDIAEF 29

RESULT 9

B71095

hypothetical protein PH1023 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: B71095

R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Hakiwa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Ogi
DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137
A:Accession: B71095

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-598 <KAW>

A:Cross-references: GB:AP000004; NID:g3236131; PIDN:BA30120.1; PID:g3257437
A:Experimental source: strain OT3
C:Genetics:

A:Gene: PH1023

Query Match
Best Local Similarity 79.4%; Score 27; DB 2; Length 598;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVWVAAF 8
11: |||
Db 280 EIVNVAEF 287

RESULT 10

D96001

Probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) [imported] - Sinorhizobium m

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C:Accession: D96001

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: D96001

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC49676.1; PID:g15141163; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Gallbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl

pele, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
heault, P.; Vandemol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:

A:Gene: glgB2; SMB21447
A:Gene: plasmid

C:Superfamily: trehalose trehalohydrolase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match
Best Local Similarity 79.4%; Score 27; DB 2; Length 601;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVWVAAF 8
11: |||
Db 151 EIMVAAF 158

RESULT 11

T05596
Probable potassium transport protein F9D16.110 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 18-Aug-2000
C:Accession: T05596
R:Bayan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hohenseil, J.; Mewes, H.W.; Meyer, K.F.
Submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15419
A:Accession: T05596
A:Molecule type: DNA
A:Residues: 1-802 <BEV>
A:Cross-references: EMBL:AL035394
A:Experimental source: cultivar Columbia; BAC clone F9D16
C:Genetics:
A:Map position: 4
A:Introns: 1/3; 16/3; 102/1; 185/1; 209/1; 296/1; 313/3; 352/1; 437/1
A:Note: F9D16.110
C:Superfamily: barley probable potassium transport protein HAK1
C:Keywords: ion transport

Query Match 79.4%; Score 27; DB 2; Length 802;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 VMXVAEF 8
||:||||
Db 629 VMSIAEF 635

RESULT 12
T30282
calcium-binding protein - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T30282
R:Solystik-Espanola, M.; Kilmzing, D.C.; Pfarr, K.; Burke, R.D.; Ernst, S.G.
Dev. Biol. 165, 73-85, 1994
A>Title: Endo16, a large multidomain protein found on the surface and ECM of endodermal
A:Reference number: Z20805; MUID:94374583
A:Accession: T30282
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1560 <SC0>
A:Cross-references: EMBL:L34680; NID:g511893; PID:g511894; PIDN:AAA30047.1
C:Genetics:
A:Note: endo16

Query Match 79.4%; Score 27; DB 2; Length 1560;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVMXVAEF 8
||:||||
Db 705 ELAAVAEF 712

RESULT 13
F95318
conserved hypothetical protein Sma0833 [imported] - Sinorhizobium meliloti (strain 1021)
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: F95318
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A>Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: F95318
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-46 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK5112.1; PID:914523550; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSyma

R:Gallibert, F.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R. L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; LeJau
habault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma0833
A:Genome: plasmid

Query Match 76.5%; Score 26; DB 2; Length 46;
Best Local Similarity 71.4%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVMXVAE 7
||:||||
Db 17 EIMAVAE 23

RESULT 14
B55209
H transfer determinant A - plasmid R27
C:Species: plasmid R27
C>Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 05-Nov-1999
C:Accession: B55209
R:Whelan, K.F.; Maher, D.; Colleran, E.; Taylor, D.E.
J. Bacteriol. 176, 2242-2251, 1994
A>Title: Genetic and nucleotide sequence analysis of the gene htdA, which regulates c
A:Reference number: A55209; MUID:94209223
A:Accession: B55209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-150 <WHE>
A:Cross-references: GB:L20342; NID:g410304; PIDN:AAB00506.1; PID:g1326032
C:Genetics:
A:Gene: htdA
A:Genome: plasmid

Query Match 76.5%; Score 26; DB 2; Length 150;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

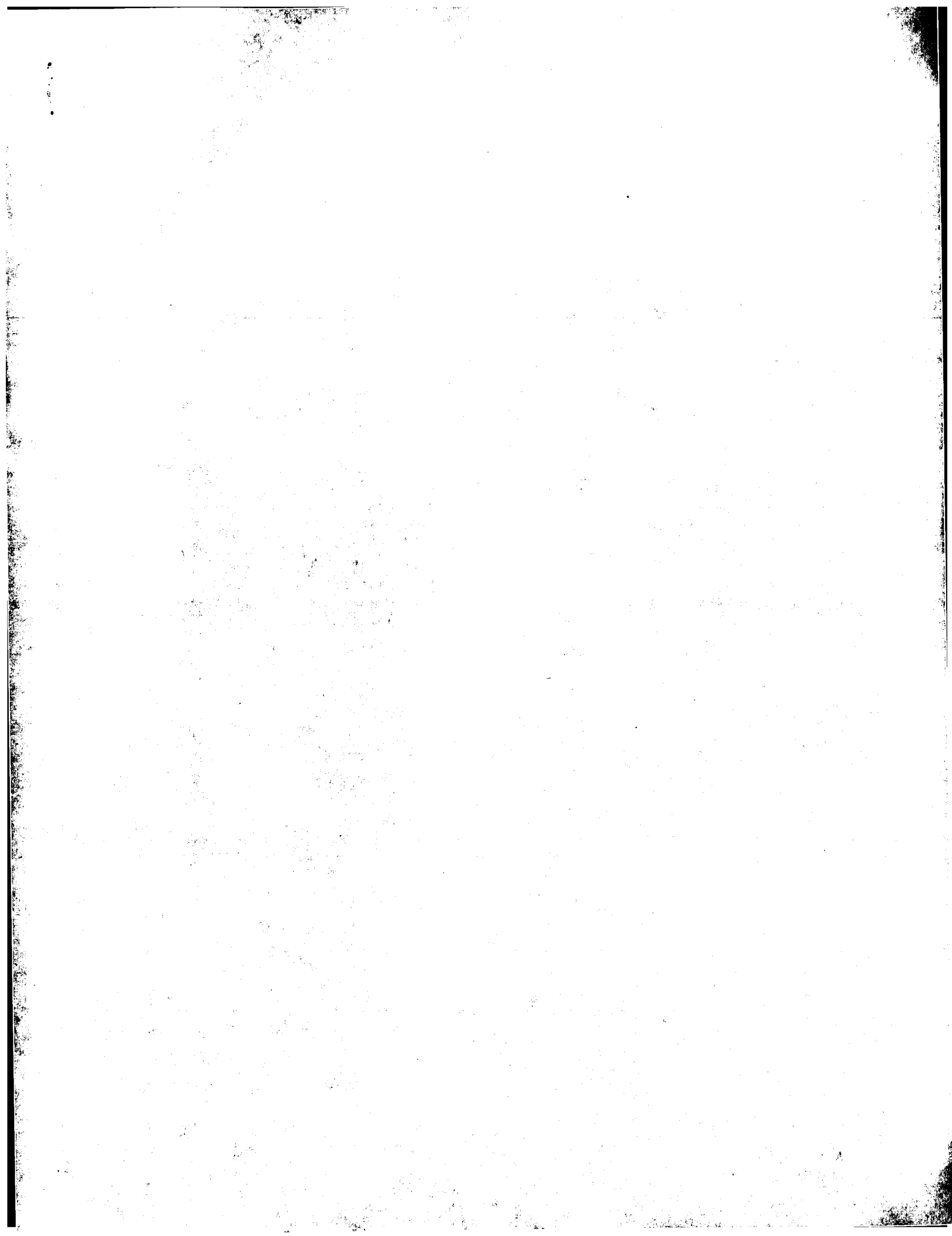
OY 1 EVMXVAEF 8
||:||||
Db 18 EVLISIEF 25

RESULT 15
E64205
deoxyribose-phosphate aldolase (EC 4.1.2.4) - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C>Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 21-Jul-2000
C:Accession: E64205
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.
M.; Fuhrman, J.; Nguyen, D.; Uitterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,
C.A.; Venter, J. C.
Science 270, 397-403, 1995
A>Title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; MUID:96026346
A:Accession: E64205
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-223 <TIG>
A:Cross-references: GB:U03684; GB:LA3967; NID:g38444650; PIDN:AACT1266.1; PID:g1045723
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: GCG3
C:Superfamily: deoxyribose-phosphate aldolase
C:Keywords: aldehyde lyase; carbon-carbon lyase

Query Match 76.58; Score 26; DB 2; Length 223;
Best Local Similarity 71.48; Pred. No. 59;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VMXVAEF 8
11 : 111
DB 93 VMNTAEF 99

Search completed: October 30, 2002, 12:31:34
Job time : 4.53317 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 30, 2002, 12:21:10 ; Search time 0.687961 Seconds
(without alignments)
450.253 Million cell updates/sec

Title: US-09-724-571-81

Perfect score: 34

Sequence: 1 EYMXVAEP 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSPROT_40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	85.3	236	1	YP58_MYCTU
2	28	82.4	3414	1	POLG_LANVT
3	27	79.4	181	1	NOO2_THFTH
4	27	79.4	253	1	YRBO_BACSU
5	26	76.5	223	1	DEOC_MYCCE
6	26	76.5	224	1	DEOC_MYCPN
7	26	76.5	269	1	SRPB_MOUSE
8	26	76.5	271	1	SRPB_HUMAN
9	26	76.5	457	1	ARLY_PASMU
10	26	76.5	603	1	SYD_AOUAE
11	25	73.5	152	1	YABE_ECOLI
12	25	73.5	413	1	ZABB_RABIT
13	25	73.5	426	1	ZABA_PIG
14	25	73.5	443	1	ZABB_HUMAN
15	25	73.5	443	1	ZABB_PIG
16	25	73.5	447	1	ZABA_HUMAN
17	25	73.5	447	1	ZABA_RAT
18	25	73.5	447	1	ZABA_RAT
19	25	73.5	453	1	ZABD_RAT
20	25	73.5	499	1	ZABA_DROME
21	25	73.5	513	1	BMP6_HUMAN
22	25	73.5	632	1	EPFD_SCHPO
23	25	73.5	808	1	RCR4_MOUSE
24	25	73.5	4466	1	DTHC_ANTCR
25	25	73.5	4466	1	DTHC_ANTCR
26	25	73.5	4486	1	DYH9_HUMAN
27	25	73.5	91	1	YAHQ_ECOLI
28	24	70.6	145	1	RL13_HALMA
29	24	70.6	148	1	CYC6_CHIRE
30	24	70.6	157	1	RAP_TAROF
31	24	70.6	186	1	RS7_METJA
32	24	70.6	187	1	REI_METJA
33	24	70.6	263	1	NIH2_METIV

34	24	70.6	291	1	AMPN_ARCFU	028438 archaeoglob
35	24	70.6	295	1	STOE_BOVIN	P19217 bos taurus
36	24	70.6	298	1	YSMK_CAEL	Q19408 caenorhabdi
37	24	70.6	341	1	KHYB_ECOLI	P00557 escherichia
38	24	70.6	377	1	NCK1_HUMAN	P16333 homo sapien
39	24	70.6	402	1	OPS4_CANAL	P46596 candida alb
40	24	70.6	431	1	PURA_EDWIC	031047 edwardsiell
41	24	70.6	474	1	YPC2_CAEL	Q11179 caenorhabdi
42	24	70.6	549	1	TCPA_CAEL	P41988 caenorhabdi
43	24	70.6	565	1	TP6B_AERPE	Q9Y64 aeropyrum p
44	24	70.6	758	1	VKGC_BOVIN	007175 bos taurus
45	24	70.6	758	1	VKGC_HUMAN	P38435 homo sapien

ALIGNMENTS

RESULT 1
ID YP58_MYCTU STANDARD: PRT: 236 AA.
AC 050740:

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 25.7 kDa protein RV2558.
GN RV2558 OR MT2635 OR MYCY9C4.10C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
[1]
SEQUENCE FROM N.A.

RA MEDLINE=96295987; PubMed=6634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Stulson J.E., Taylor K., Whitehead S., Barrrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
[2]
SEQUENCE FROM N.A.

RC STRAIN=CD 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
-i- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV2557.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z77250; CAB01046.1; -;
DR EMBL: AE007098; AAK46947.1; -;
DR TIGR: MT2635; -;
DR Tuberculist: RV2558; -;
KW Hypothetical protein: Complete proteome.
SO SEQUENCE 236 AA; 25718 MW; 13E3B049DBF79C6B CRC64;

Query Match 85.3%; Score 29; DB 1; Length 236;
 Best Local Similarity 75.0%; Pred. No. 4.5;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVMXAEF 8
 11:1111
 Db 216 EVDVAEF 223

RESULT 2

POLG_LANVT STANDARD; PRT; 3414 AA.
 ID POLG_LANVT STANDARD; PRT; 3414 AA.
 AC P29837;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polypeptide protein C (core protein);
 DE Envelope protein E; Nonstructural protein M; Major
 DE NS2A; Nonstructural protein NS2B; Helicase/protease (EC 3.4.21.98)
 DE (NS3); Nonstructural protein NS4A; Nonstructural protein NS4B; RNA-
 DE directed RNA polymerase (EC 2.7.7.48) (NS5).
 OS Langkat virus (strain TP21).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 NCBI_TaxID=31638;
 RX MEDLINE=92263794; PubMed=1720591;
 RA Mandl C.W., Iacono-Connors L., Wallner G., Holzmann H., Kunz C.,
 RA Heinz F.X.;
 RT "Sequence of the genes encoding the structural proteins of the low-
 RT virulence tick-borne flaviviruses Langkat TP21 and Yelantsev.";
 RL Virology 185:891-895(1991).
 RN [2]
 RP SEQUENCE OF 777-3414 FROM N.A.
 RP MEDLINE=92263794; PubMed=1316684;
 RA Iacono-Connors L.C., Schmaljohn C.S.;
 RT "Cloning and sequence analysis of the genes encoding the
 RT nonstructural proteins of Langkat virus and comparative analysis with
 RT other flaviviruses.";
 RL Virology 188:875-880(1992).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF FOUR PEPTIDE BONDS IN THE VIRAL
 CC PRECURSOR POLYPEPTIDE, COMMONLY WITH ASP OR GLU IN THE P6
 CC POSITION, CYS OR THR IN P1 AND SER OR ALA IN P1'.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPID-PROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND RNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M73835; AAA02740.1; ALT_TERM.
 CC EMBL: S35365; AAB22165.1; -.
 DR PIR: A41704; A41704.
 DR PIR: A42545; A42545.
 DR HSSP: P14336; 1SVB.
 DR MEROPS: S07.001; -.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR000069; Flav1_M.
 DR InterPro: IPR001157; Flav1_NS1.
 DR InterPro: IPR000752; Flav1_NS2A.
 DR InterPro: IPR000487; Flav1_NS2B.
 DR InterPro: IPR000404; Flav1_NS4A.
 DR InterPro: IPR001528; Flav1_NS4B.
 DR InterPro: IPR000208; Flav1_NS5.
 DR InterPro: IPR001122; Flav1_capsid.

DR InterPro: IPR000336; Flav1_glycoprote.
 DR InterPro: IPR001850; Flav1_helicase.
 DR InterPro: IPR002535; Flav1_Prop.
 DR InterPro: IPR002877; FtsJ.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF01003; Flav1_capsid; 1.
 DR Pfam: PF00869; Flav1_glycoprot; 1.
 DR Pfam: PF02832; Flav1_glycop_C; 1.
 DR Pfam: PF00949; Flav1_helicase; 1.
 DR Pfam: PF01004; Flav1_M; 1.
 DR Pfam: PF00948; Flav1_NS1; 1.
 DR Pfam: PF01005; Flav1_NS2A; 1.
 DR Pfam: PF01002; Flav1_NS2B; 1.
 DR Pfam: PF01350; Flav1_NS4A; 1.
 DR Pfam: PF01349; Flav1_NS4B; 1.
 DR Pfam: PF00972; Flav1_NS5; 1.
 DR Pfam: PF01570; Flav1_Prop; 1.
 DR Pfam: PF01728; FtsJ; 1.
 DR ProDom: PD001496; Flav1_NS1; 1.
 DR ProDom: PD001556; Flav1_glycoprote; 1.
 DR SMART: SM00490; HELIC_C; 1.
 KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
 KW Core protein; Coat protein; Envelope protein; Hydrolyase; Helicase;
 KW ATP-binding; Transmembrane; Nonstructural protein.
 FT INIT_MET 1
 FT CHAIN 1
 FT CHAIN 113 112
 FT CHAIN 206 280
 FT CHAIN 281 776
 FT CHAIN 777 1128
 FT CHAIN 1129 1358
 FT CHAIN 1359 1489
 FT CHAIN 1490 2110
 FT CHAIN 2111 2259
 FT CHAIN 2260 2511
 FT CHAIN 2512 3414
 FT NP_BIND 1688 1695
 FT SITE 1779 1782
 FT TRANSMEM 103 119
 FT TRANSMEM 262 278
 FT TRANSMEM 728 744
 FT TRANSMEM 758 774
 FT DISULFID 283 310
 FT DISULFID 340 396
 FT DISULFID 354 385
 FT DISULFID 372 401
 FT DISULFID 466 570
 FT DISULFID 587 618
 FT CARBOHYD 144 144
 FT CARBOHYD 434 434
 SO SEQUENCE 3414 AA; 378017 MW; 59CB7E95DD70B82E CRC64;
 Query Match 82.4%; Score 28; DB 1; Length 3414;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 VMXAEF 8
 111111
 Db 929 VMTVAEF 935

RESULT 3
 NOO2_THETH STANDARD; PRT; 181 AA.
 ID NOO2_THETH STANDARD; PRT; 181 AA.
 AC O56221;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 2 (EC 1.6.5.3) (NADH
 dehydrogenase 1, chain 2) (NDH-1, chain 2).

```

GN  N002.
OS  Thermus aquaticus (subsp. thermophilus).
CC  Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX  NCBI_TaxID=274;
RN  [1]
RP  SEQUENCE FROM N.A. AND CHARACTERIZATION.
RC  STRAIN-HB8 / ATCC 27634;
RX  MEDLINE=97172490; PubMed=9020134;
RA  Yano T., Chu S.S., Sled V.D., Ohnishi T., Yagi T.;
RT  "The proton-translocating NADH-quinone oxidoreductase (NDH-1) of
RT  thermophilic bacterium Thermus thermophilus HB-8. Complete DNA
RT  sequence of the gene cluster and thermostable properties of the
RT  expressed N002 subunit."
RL  J. Biol. Chem. 272:4201-4211(1997).
CC  -1- CATALYTIC ACTIVITY: NADH + ubiquinol = NAD(+) + ubiquinol.
CC  -1- COFACTOR: BINDS A 2FE-2S CLUSTER.
CC  -1- SUBUNIT: COMPOSED OF 14 DIFFERENT SUBUNITS.
CC  -1- SIMILARITY: BELONGS TO THE COMPLEX I 24 KDA SUBUNIT FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement. (See http://www.isb-sdb.ch/announce/
CC  or send an email to license@sdb.ch.)
CC  -----
DR  EMBL: U52917; AAA97942.1; -
DR  InterPro: IPR002023; Complex1_24KD.
DR  Pfam: PF01257; Complex1_24KD; 1.
DR  Prodom: PD003859; Complex1_24KD; 1.
DR  ProSITE: PS01099; Complex1_24K; 1.
KM  Oxidoreductase; NAD; Ubiquinol; Iron-sulfur.
FT  METAL 83 83
FT  METAL 88 88
FT  METAL 124 124
FT  METAL 128 128
SQ  SEQUENCE 181 AA; 20286 MW; 484FE09245C613EE CRC64;

Query Match
Best Local Similarity 79.4%; Score 27; DB 1; Length 181;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY  1 EVXVXAEF 8
DB  59 EVMGVASF 66

RESULT 4
YTBQ_BACSU
ID  YTBQ_BACSU STANDARD: PRT; 253 AA.
AC  P53560;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  01-MAR-2002 (Rel. 41, Last annotation update)
DE  Hypothetical protein ytbQ.
GN  YTBQ.
OS  Bacillus subtilis.
CC  Bacteria; Firmicutes; Bacillus/Clostridium group;
CC  Bacillus/Staphylococcus group; Bacillus.
OX  NCBI_TaxID=1423;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=96312354; PubMed=8763940;
RA  Bower S., Perkins J.B., Vocum R.R., Howitt C.L., Rahaim P.,
RA  Pero J.;
RT  "Cloning, sequencing, and characterization of the Bacillus subtilis
RT  biotin biosynthetic operon."
RL  J. Bacteriol. 178:4122-4130(1996).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its

```

```

CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement. (See http://www.isb-sdb.ch/announce/
CC  or send an email to license@sdb.ch.)
CC  -----
DR  EMBL: U51868; AAB17463.1; -
DR  Subtilisin; BG11787; ytbQ.
KM  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 253 AA; 28233 MW; 572CC4681F04C650 CRC64;

Query Match
Best Local Similarity 79.4%; Score 27; DB 1; Length 253;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY  1 EVXVXAEF 8
DB  74 DIMDIAEF 81

RESULT 5
DEOC_MYCGE
ID  DEOC_MYCGE STANDARD: PRT; 223 AA.
AC  P47296;
DT  01-FEB-1996 (Rel. 33, Created)
DT  01-FEB-1996 (Rel. 33, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Deoxyribose-phosphate aldolase (EC 4.1.2.4) (Phosphodeoxyribosaldolase)
DE  (Deoxyribosaldolase).
GN  DEOC OR MG050.
OS  Mycoplasma genitalium.
CC  Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
CC  Mycoplasmales; Mycoplasma.
OX  NCBI_TaxID=2097;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=ATCC 33530 / G-37;
RX  MEDLINE=96026346; PubMed=756993;
RA  Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA  Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA  Nguyen D.T., Ullrich T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA  Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Luetter T.S.,
RA  Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT  "The minimal gene complement of Mycoplasma genitalium."
RL  Science 270:397-403(1995).
CC  -1- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 5-phosphate = D-
CC  glyceraldehyde 3-phosphate + acetaldehyde.
CC  -1- PATHWAY: NUCLEOTIDE AND DEOXYRIBONUCLEOTIDE CATABOLISM.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -1- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES.
CC  DEOC SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement. (See http://www.isb-sdb.ch/announce/
CC  or send an email to license@sdb.ch.)
CC  -----
DR  EMBL: U39684; AAC71266.1; -
DR  TIGR: MG050; -
DR  InterPro: IPR002915; Deoc.
DR  InterPro: IPR003009; FMN_enzyme.
DR  Pfam: PF01791; Deoc; 1.
KM  Lyase; Schiff base; Complete proteome.
FT  BINDING 152 152 SCHIFF-BASE (BY SIMILARITY).
SQ  SEQUENCE 223 AA; 24675 MW; 332430231CE99DB0 CRC64;

Query Match
Best Local Similarity 76.5%; Score 26; DB 1; Length 223;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

OY 2 VMXVAEF 8
11 :|||
DB 93 VMNIAEF 99

```

RESULT 6
DEOC MYCPN STANDARD; PRT; 224 AA.
ID DEOC MYCPN
AC P09924;
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Deoxyribose-phosphate aldolase (EC 4.1.1.2.4) (Phosphodeoxyribosomalolase)
DE (Deoxyribosomalolase).
GN DEOC OR MPN053 OR MP091.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RA MEDLINE=89128453; PubMed=2492658;
RA Loebel S., Inamine J.M., Hu P.-C.;
RA "Nucleotide sequence of the deoc gene of Mycoplasma pneumoniae.";
RL Nucleic Acids Res. 17:801-801(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RA MEDLINE=97105885; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Hermann R.;
RA "Complete sequence analysis of the genome of the bacterium Mycoplasma
RA pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 5-phosphate = D-
CC glycerolaldehyde 3-phosphate + acetaldehyde.
CC -1- PATHWAY: NUCLEOTIDE AND DEOXYRIBONUCLEOTIDE CATABOLISM.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES.
CC DEOC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X13544; CAA31897.1; -.
DR EMBL: AE000011; AAB95739.1; -.
DR PIR: S02216; S02216.
DR InterPro: IPR002915; Deoc.
DR Pfam: PF01791; Deoc. 1.
KW Lyase; Schiff base; Complete proteome.
FT BINDING 152 152 SCHIFF-BASE (BY SIMILARITY).
FT SEQUENCE 224 AA; 24878 MW; 73C3E4932E7881F7 CRC64;
SO
Query Match 76.5%; Score 26; DB 1; Length 224;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Signal recognition particle receptor beta subunit (SR-beta).
DE SRPB.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95146535; PubMed=7844142;
RX Miller J.D., Tajima S., Lauffer L., Walter P.;
RT "The beta subunit of the signal recognition particle receptor is a
RT transmembrane GTPase that anchors the alpha subunit, a peripheral
RT membrane GTPase, to the endoplasmic reticulum membrane.";
RL J Cell Biol. 128:273-282(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Small intestine;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arawaka T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
RA Blake J., Boffelli D., Bojunga N., Carinici P., de Bonaldo M.F.,
RA Brownstein M.J., Bolt C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: HAS GTPASE ACTIVITY. MAY MEDIANE THE MEMBRANE
CC ASSOCIATION OF SR ALPHA.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane.
CC -1- SIMILARITY: TO OTHER SRP BETA SUBUNITS; DISTANTLY RELATED TO RAS
CC SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U17343; AAB69976.1; -.
DR EMBL: AK008383; BAB25638.1; -.
DR MGD: MGI:102964; SrpB.
DR InterPro: IPR003575; Small_GTPase.
DR SMART: SM00010; small_GTPase; 1.
KW Signal recognition particle; Transmembrane; Receptor;
KW Endoplasmic reticulum; GTP-binding.
FT TRANSMEM 35 55 POTENTIAL.
FT NP_BIND 69 76 GTP (POTENTIAL).
FT NP_BIND 115 119 GTP (POTENTIAL).
FT NP_BIND 178 181 GTP (POTENTIAL).
FT CONFLICT 16 16 A -> P (IN REF. 2).
FT CONFLICT 173 173 L -> P (IN REF. 1; AA SEQUENCE).
FT SEQUENCE 269 AA; 29579 MW; 041175FA6891DA37 CRC64;
SO
Query Match 76.5%; Score 26; DB 1; Length 269;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```



```

RT aeolicus."
RL Nature 392:353-358(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspareryl-tRNA(Asp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000750; AAC07548.1; -.
DR HSSP: P36419; IEFW.
DR InterPro: IPR002106; AA_tRNA_ligase_II.
DR InterPro: IPR004115; GAD.
DR InterPro: IPR002309; tRNA-synt_2.
DR InterPro: IPR002312; tRNA-synt_asep.
DR Pfam: PF02938; GAD; 1.
DR Pfam: PF00152; tRNA-synt_2; 2.
DR Pfam: PF01336; tRNA-anti; 1.
DR PRINTS: PR01042; TRNASYNTHASP.
DR PROSITE: PS00179; AA_TRNA_LIGASE_II_1; 1.
DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; 1.
DR KW Aminoacyl-tRNA synthetase; protein biosynthesis; ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 603 AA; 69729 MW; 9DFFBD840C8DC1C CRC64;

Query Match 76.5%; Score 26; DB 1; Length 603;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVMXVAE 7
   ||| |||
Db 254 EVMQVAE 260

RESULT 11
YABB.ECOLI STANDARD; PRT; 152 AA.
ID YABB.ECOLI
AC P22186;
DC 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yabb.
GN YABB OR B0081.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN NCBI_TaxID=562;
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=90251464; PubMed=2187182;
RA Gomez M.J., Flouret B., van Heijenoort J., Ayala J.A.;
RT "Nucleotide sequence of the regulatory region of the gene pppB of
RT Escherichia coli."
RL Nucleic Acids Res. 18:2813-2813(1990).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- SIMILARITY: BELONGS TO THE UPF0040 FAMILY. STRONG, TO H.INFLUENZAE

```

```

CC H1129.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52063; CA36284.1; -.
DR EMBL: X55034; CA38858.1; -.
DR EMBL: AE000118; AAC73192.1; -.
DR PIR: S14388; S14388.
DR Ecogene: EG11084; Yabb.
DR InterPro: IPR003444; UPF0040.
DR Pfam: PF02381; UPF0040; 2.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 152 AA; 17386 MW; 3EE1A6FA9D2B1C01 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 152;
Best Local Similarity 62.5%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVMXVAEF 8
   ||| | :|
Db 107 EVMVGQF 114

RESULT 12
ZABB.RABIT STANDARD; PRT; 413 AA.
ID ZABB.RABIT
AC Q00006;
DC 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
DE beta isoform (PP2A, subunit B, B-beta isoform) (PP2A, subunit B,
DE R2-beta isoform) (PP2A, subunit B, PR55-beta isoform) (PP2A, subunit B,
DE PP2R2B.
GN Oryctolagus cuniculus (Rabbit).
OS Oryctolagus cuniculus; Chordata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN NCBI_TaxID=9986;
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=91198016; PubMed=1849734;
RA Mayer R.E., Hendrix P., Cron P., Mathies R., Stone S.R.,
RA Goris J., Merlevede W., Horstenge J., Hemmings B.A.;
RT "Structure of the 55-kDa regulatory subunit of protein phosphatase
RT 2A: evidence for a neuronal-specific isoform."
RL Biochemistry 30:3589-3597(1991).
CC -1- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
CC COMPARTMENT.
CC -1- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATE
CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
CC B (THE R2/B/PR55/B55, R3/B-/PR72/PR130/PR59 AND R5/B-/B56
CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
CC AND CELL SIGNALING MOLECULES.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M64931; AAA31458.1; -
 DR InterPro: IPR000009; PP2A_PR55.
 DR Pfam: PF00400; WD40; 5.
 DR SMART: SM00320; WD40; 2.
 DR PROSITE: PS01024; PR55_1; 1.
 DR PROSITE: PS01025; PR55_2; 1.
 DR MultiGene family.
 FT NON_TER 1.
 SQ SEQUENCE 413 AA; 48243 MW; 4923787817EB8FE2 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 413;
 Best Local Similarity 62.5%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVKXVAEF 8
 ||: |||
 Db 193 EVITAAEF 200

RESULT 13
 2ABA_PIG STANDARD; PRT; 426 AA.
 AC Q29050;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
 DE alpha isoform (PP2A, subunit B, B-alpha isoform) (PP2A, subunit B,
 DE B55-alpha isoform) (PP2A, subunit B, PR55-alpha isoform) (PP2A,
 DE subunit B, R2-alpha isoform) (Fragment).
 GN PP2A2A.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NC NCBL_TaxID=9823;
 RX [1]
 RA SEQUENCE FROM N.A.
 RA Mayer-Jaekel R.E.;
 RL Thesis (1992), Friedrich Miescher Institut / Basel, Switzerland.
 CC -1- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
 CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
 CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
 CC COMPARTMENT.
 CC -1- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
 CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
 CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES
 CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
 CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
 CC B (THE R2/B/PR55/B55, R3/B'/PR72/PR130/PR59 AND R5/B'/B56
 CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
 CC AND CELL SIGNALING MOLECULES.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z34932; GA84404.1; -
 DR InterPro: IPR000009; PP2A_PR55.
 DR Pfam: PF00400; WD40; 5.
 DR PRINTS: PR00600; PP2APR55.
 DR SMART: SM00320; WD40; 3.

DR PROSITE: PS01024; PR55_1; 1.
 DR PROSITE: PS01025; PR55_2; 1.
 DR MultiGene family.
 FT NON_TER 1.
 SQ SEQUENCE 426 AA; 49613 MW; 3AAND7EB38B03534 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 426;
 Best Local Similarity 62.5%; Pred. No. 82;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVKXVAEF 8
 ||: |||
 Db 206 EVITAAEF 213

RESULT 14
 2ABH_HUMAN STANDARD; PRT; 443 AA.
 AC Q00005;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
 DE beta isoform (PP2A, subunit B, B-beta isoform) (PP2A, subunit B,
 DE beta isoform) (PP2A, subunit B, PR55-beta isoform) (PP2A, subunit B,
 DE R2-beta isoform).
 GN PP2AR2B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCBL_TaxID=9606;
 RX [1]
 RA SEQUENCE FROM N.A.
 RA Tissue-Fetal brain;
 RX MEDLINE=91196016; PubMed=1849734;
 RA Mayer R.E., Hendrix P., Cron P., Mathies R., Stone S.R.,
 RA Goyis J., Merlevede W., Hofsteenge J., Hemmings B.A.;
 RT 2A: evidence for a neuronal-specific isoform.;
 RL Biochemistry 30:3589-3597(1991).
 CC -1- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
 CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
 CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
 CC COMPARTMENT.
 CC -1- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
 CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
 CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES
 CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
 CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
 CC B (THE R2/B/PR55/B55, R3/B'/PR72/PR130/PR59 AND R5/B'/B56
 CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
 CC AND CELL SIGNALING MOLECULES.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M64930; AAA36493.1; -
 DR PIR: B38351; B38351.
 DR MIM: 604325; -
 DR InterPro: IPR000009; PP2A_PR55.
 DR Pfam: PF00400; WD40; 6.
 DR PRINTS: PR00600; PP2APR55.
 DR SMART: SM00320; WD40; 3.
 DR PROSITE: PS01024; PR55_1; 1.
 DR PROSITE: PS01025; PR55_2; 1.

KW Multigene family.
SQ SEQUENCE 443 AA; 51710 MW; C383C834B2852B8F CRC64;

Query Match 73.5%; Score 25; DB 1; Length 443;
Best Local Similarity 62.5%; Pred. No. 86;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Search completed: October 30, 2002, 12:27:52
Job time : 2.68796 secs

OY 1 EVMXVAEF 8
||: |||
DB 223 EVITAAEF 230

RESULT 15

2ABB_PIG STANDARD; PRT; 443 AA.
ID 2ABB_PIG
AC P54614;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
beta isoform (PP2A, subunit B, B-beta isoform) (PP2A, subunit B, B55-
beta isoform) (PP2A, subunit B, PR55-beta isoform) (PP2A, subunit B,
R2-beta isoform).
GN PP2R2B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_Taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Mayer-Jaekel R.E.;
RL Thesis (1992), Friedrich Miescher Institut / Basel, Switzerland.
CC -I- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
COMPARTMENT.
CC -I- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES
WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
B (THE R2/B/PR55/B55, R3/B'/PR72/PR130/PR59 AND R5/B'/B56
FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
AND CELL SIGNALING MOLECULES.
CC -I- TISSUE SPECIFICITY: BRAIN.
CC -I- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Z34933; CAA84405.1;
DR InterPro: IPR000009; PP2A_PR55.
DR Pfam: PF00400; WD40; 5.
DR PRINTS: PRO0600; PP2APR55.
DR SMART: SM00320; WD40; 4.
DR PROSITE: PS01024; PR55_1; 1.
DR PROSITE: PS01025; PR55_2; 1.
KW Multigene family.
SQ SEQUENCE 443 AA; 51459 MW; F8562FC696719F41 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 443;
Best Local Similarity 62.5%; Pred. No. 86;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 EVMXVAEF 8
||: |||
DB 223 EVITAAEF 230

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:20:50 ; Search time 1.60197 seconds

(without alignments)
277.344 Million cell updates/sec

Title: US-09-724-571-103

Perfect score: 20

Sequence: 1 VKMD 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_032802:*

```
1: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/genesqp-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	4	21	AA832122
2	20	100.0	4	21	AA87945
3	20	100.0	5	18	AA08216
4	20	100.0	5	19	AA61150
5	20	100.0	5	20	AA33750
6	20	100.0	5	22	AA87260
7	20	100.0	8	14	AA42401
8	20	100.0	8	21	AA94772
9	20	100.0	8	22	AAE10660
10	20	100.0	8	22	AAE06902
11	20	100.0	8	22	AAU06631

12	20	100.0	8	22	AAU06635	Synthetic fluoresc
13	20	100.0	8	22	AAU07230	Human beta-amyloid
14	20	100.0	8	22	AAE02612	Human Aspartyl pro
15	20	100.0	9	19	AA82084	Fluorogenic protea
16	20	100.0	9	19	AA82083	Fluorogenic protea
17	20	100.0	9	21	AA807873	A peptide fragment
18	20	100.0	9	21	AA87949	Mammalian amyloid
19	20	100.0	10	13	AA82054	Peptide p1. Synth
20	20	100.0	10	13	AA82461	Human amyloid pr
21	20	100.0	10	20	AA82440	Human amyloid beta
22	20	100.0	10	21	AA169703	Beta-Ap alpha-sec
23	20	100.0	10	22	AAE10654	Human wild-type AP
24	20	100.0	10	22	AAE06899	Human amyloid prec
25	20	100.0	10	22	AAU06628	ASP2 recognition s
26	20	100.0	10	22	AAU07227	Human beta-amyloid
27	20	100.0	10	22	AA662668	Beta-sheet breaker
28	20	100.0	10	22	AAE02606	Human wild-type AP
29	20	100.0	10	22	AA866574	Synthetic peptide
30	20	100.0	10	22	AA846205	Human APP derived
31	20	100.0	10	22	AA846206	Human APP derived
32	20	100.0	10	22	AA846207	Human APP derived
33	20	100.0	10	22	AA846208	Human APP derived
34	20	100.0	10	22	AA846209	Human APP derived
35	20	100.0	10	22	AA846210	Human APP derived
36	20	100.0	10	22	AA846213	Synthetic peptide f
37	20	100.0	11	22	AA875143	ASP 1 substrate se
38	20	100.0	11	22	AA875144	ASP2 substrate wil
39	20	100.0	11	22	AA874468	Beta-amyloid precu
40	20	100.0	12	22	AA874931	Quenched fluoresce
41	20	100.0	13	19	AA870869	Human Asp-2b activ
42	20	100.0	15	22	AAE10669	Human beta-amyloid
43	20	100.0	15	22	AAE06906	Human beta-amyloid
44	20	100.0	15	22	AAU07234	Human beta-amyloid
45	20	100.0	16	21	AA806315	Human beta-amyloid

ALIGNMENTS

RESULT 1	
AA832122	
ID	AA832122 standard; peptide: 4 AA.
AC	AA832122;
XX	
DE	14-FEB-2001 (first entry)
XX	
DE	Beta-secretase target region of amyloid-beta precursor protein.
XX	
KW	Aldehyde caspase inhibitor; antibody; amyloid-beta precursor;
KW	caspase cleavage; neurodegenerative disease.
XX	
OS	Synthetic.
XX	
PN	WO200063250-A1.
XX	
PD	26-OCT-2000.
XX	
PF	13-APR-2000; 2000WO-CA00414.
XX	
PR	15-APR-1999; 99US-0129495.
XX	
PA	(MERI) MERCK PROSST CANADA & CO.
XX	
PI	Gervais F, Roy S, Nicholson DW, Xu D, Robertson G, Huang J;
XX	
DR	WPI: 2000-687160/67.
XX	
PT	Novel antibody that recognizes neo-epitopes of caspase cleaved
PT	amyloid-beta precursor protein useful for diagnosing conditions
PT	involving neuronal apoptosis such as Alzheimer's, Huntington's and
XX	Parkinson's disease -
XX	

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VKMD 4
 ||||
 Db 1 VKMD 4

RESULT 4
 AAM61150
 ID AAM61150 standard; Peptide: 5 AA.
 XX AAM61150;
 AC AAM61150;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Wild-type APP cleavage site.
 XX
 KM Beta-secretase; human; beta-amyloid precursor protein; APP;
 KM Protease; Inhibitor; screening; Alzheimer's disease; therapy.
 XX
 OS Homo sapiens.
 XX
 PN W09826059-A1.
 XX
 PD 18-JUN-1998.
 XX
 PF 11-DEC-1996; 96WO-US19549.
 XX
 PR 11-DEC-1996; 96WO-US19549.
 XX
 PA (ATHE-) ATHENA NEUROSCIENCES INC.
 XX
 PI Anderson JP, Chrysler SMS, Kelm PS, Sinha S;
 XX
 DR WPI: 1998-348519/30.
 XX
 PR Novel beta-secretase which cleaves beta-amyloid precursor protein -
 PT useful for screening for compounds which inhibit the cleavage and
 PS are useful for treating Alzheimer's disease
 XX
 XX Disclosure: Page 20; 39pp; English.

CC This peptide comprises the site of wild-type beta-amyloid precursor
 CC protein (APP) (MBP-C125 WT) that is cleaved by a novel
 CC beta-secretase isolated from human 293 cells. This protease
 CC cleaves APP at the N-terminus of the beta-amyloid peptide (beta-AP).
 CC It is believed to be the putative beta-secretase responsible for
 CC the pathogenic processing of APP to beta-AP in Alzheimer's disease,
 CC Down's syndrome and HCHWA-D. Recombinant fusion proteins were
 CC generated comprising the last 125 amino acids of APP (wild-type
 CC or Swedish double mutation, see AAM61151-52) fused to the C-terminal
 CC end of maltose binding protein. The 2 fusion proteins were
 CC expressed in *Escherichia coli*, and used as substrates for
 CC beta-secretase in beta-secretase inhibitor assays. Compounds that
 CC inhibit APP cleavage by beta-secretase may be useful in the
 CC treatment of Alzheimer's disease.
 CC
 XX Sequence 5 AA:

Query Match 100.0%; Score 20; DB 19; Length 5;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VKMD 4
 ||||
 Db 1 VKMD 4

RESULT 5
 AAY33750
 ID AAY33750 standard; Protein: 5 AA.
 XX AAY33750;
 AC AAY33750;

XX 09-NOV-1999 (first entry)
 DT
 XX
 DE Wild type beta-amyloid protein precursor (APP) cleavage site.
 DE
 XX Beta-secretase; beta-amyloid protein precursor; APP; Down's syndrome;
 KM Alzheimer's disease; cleavage site; wild type.
 KM
 XX Homo sapiens.
 OS
 XX US5942400-A.
 PN
 XX
 PD 24-AUG-1999.
 XX
 PF 07-JUN-1996; 96US-0659984.
 XX
 PR 07-JUN-1996; 96US-0659984.
 PR 07-JUN-1995; 95US-0480498.
 PR 07-JUN-1995; 95US-0485152.
 XX
 PA (ELAN-) ELAN PHARM INC.
 XX
 PI Anderson JP, Jacobson-Croak KL, Sinha S;
 XX
 DR WPI: 1999-517417/43.
 XX
 PR A method for detecting human beta-secretase cleavage of polypeptides
 PT useful for identifying beta-secretase inhibitors
 PS
 XX Examples; Column 28; 43pp; English.

CC This sequence is the wild type beta-amyloid protein precursor (APP)
 CC cleavage site. APP is cleaved by beta-secretase AAY33741. The wild type
 CC cleavage site and the Swedish mutant version of this site AAY33751 are
 CC used in a method for detecting human beta-secretase cleavage of
 CC polypeptides and for identifying beta-secretase inhibitors. Inhibition
 CC of beta-secretase activity would be useful for chemical modelling of a
 CC critical event in the pathology of Alzheimer's disease. Inhibitors of
 CC beta-secretase would be useful for the prevention and treatment of
 CC Alzheimer's disease and Down's Syndrome.
 CC
 XX Sequence 5 AA:

Query Match 100.0%; Score 20; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VKMD 4
 ||||
 Db 1 VKMD 4

RESULT 6
 AAB47260
 ID AAB47260 standard; Peptide: 5 AA.
 XX AAB47260;
 AC AAB47260;
 XX
 DT 18-JUL-2001 (first entry)
 XX
 DE Wild type APP sequence for cleavage by beta-secretase.
 XX
 KM Beta-secretase; isotype; beta-amyloid precursor protein; APP;
 KM beta-amyloid peptide; beta-AP; Alzheimer's disease; Down's syndrome;
 KM HCHWA-D; Swedish mutation; maltose binding protein; MBP.
 XX
 OS Homo sapiens.
 XX
 PN US6221645-B1.
 XX
 PD 24-APR-2001.
 XX
 PF 07-JUN-1996; 96US-0660531.

```

XX 07-JUN-1995; 950S-0480498.
PR (ELAN-) ELAN PHARM INC.
PA
XX Chryslers SMS, Sinha S, Keim PS, Anderson JP, Tan H, McConlogue LC;
PI
XX WPI; 2001-315578/33.
DR
XX Novel antibody that specifically binds native beta-secretase protein,
PT useful for raising anti-idiotypic antibodies and for detecting or
PT diagnosing pathological conditions related to presence of respective
PT antigens
XX Example; Column 28; 42pp; English.
PS
XX The sequences given in AAB47260-61 represent cleavage sites derived
CC from wild-type and the Swedish mutation of beta-amyloid precursor
CC protein (APP). These cleavage sites were used in fusion proteins
CC which were used as substrates for the beta-secretase protein which
CC is characterized by an ability to cleave the 695-amino acid isoform
CC of APP between amino acids 596 and 597. The fusion proteins contain
CC the carboxy-terminal end of Maltose binding protein (MBP) fused to
CC the carboxy-terminal 125 amino acids of either wild type APP or APP
CC containing the Swedish mutation. Beta-secretase is thought to be
CC responsible for the pathogenic processing of APP to form beta amyloid
CC peptide (beta-AP) in beta-AP related conditions, e.g. Alzheimer's
CC disease, Down's syndrome, HCHWA-D etc. Beta-secretase has a molecular
CC weight of 260-300 kD and will bind to wheat germ agglutinin but not to
CC concanavalin A. Beta-secretase will cleave both the wild type and
CC the Swedish mutation of APP.
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 20; DB 22; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VKMD 4
DB 1 VKMD 4
RESULT 7
AAB42401
ID AAB42401 standard; peptide; 8 AA.
XX
AC AAB42401;
XX
DT 01-APR-1994 (first entry)
XX
DE Peptide representing a variation of the C-terminus of ATF-beta APP.
XX
KW Amyloid precursor protein; Alzheimer's disease; fragments; inhibitors;
KW amino-terminal fragment; diagnosis; screening; chemiluminescence assay.
XX
OS Synthetic.
XX
PN WO9321526-A.
XX
PD 28-OCT-1993.
XX
PF 03-MAR-1993; 93WO-US01817.
XX
PR 15-APR-1992; 92US-0868949.
XX
PR 26-OCT-1992; 92US-0965971.
XX
PA (ATHE-) ATHENA NEUROSCIENCES INC.
XX (ELIT) LILLY & CO ELI.
XX
PI Fritz LC, Schenk DB, Seubert PA;
XX
DR WPI; 1993-351873/44.

```

```

XX Monitoring beta amyloid precursor protein processing - involves
PT detecting soluble fragments from cleavage at amino terminals of
PT peptide, used to study Alzheimer's disease and potential drugs
PT for it
XX
XX Disclosure; Page 7; 38pp; English.
PS
XX The peptide represents a variant of the C-terminus of the beta-
CC amyloid precursor protein amino terminal fragment. Such a peptide
CC was used in conjunction with antibody 92 in a chemiluminescence assay
CC to probe human lumbar cerebrospinal fluid to detect Alzheimer's
CC disease. Such an assay can be used to diagnose or monitor amyloid-
CC related diseases in a patient and to screen and evaluate potential
CC drugs for the treatment of these diseases.
CC See also AAB42398-404.
XX
SQ Sequence 8 AA;
Query Match 100.0%; Score 20; DB 14; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VKMD 4
DB 5 VKMD 8
RESULT 8
AAY94772
ID AAY94772 standard; Protein; 8 AA.
XX
AC AAY94772;
XX
DT 12-FEB-2001 (first entry)
XX
DE Beta-secretase substrate peptide SEQ ID 18.
XX
KW Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease;
KW Down's syndrome; amyloid angiopathy; gene therapy; neuroprotective.
XX
OS Synthetic.
XX
PN WO200058479-A1.
XX
PD 05-OCT-2000.
XX
PF 23-MAR-2000; 2000WO-US07755.
XX
PR 26-MAR-1999; 99US-0277229.
XX
PA (AMGE-) AMGEN INC.
XX
PI Citron M, Vassar RJ, Bennett BD;
XX
DR WPI; 2000-594643/56.
XX
PT Isolated beta-secretase nucleic acids and encoded polypeptides, useful
PT for diagnosis and gene therapy of Alzheimer's disease -
XX
XX Example 10; Page 117; 145pp; English.
PS
XX This invention relates to 3 nucleotide sequences encoding beta-secretase
CC proteins. Beta-secretase is an enzyme involved in the production of one
CC of the components of amyloid plaques involved in Alzheimer's disease. The
CC invention includes an expression vector comprising the nucleotide
CC sequence, a host cell comprising the expression vector, and a process for
CC producing the protein through culturing the transformed cells. Also
CC included in the invention are a polypeptide derivative of the
CC beta-secretase protein, a fusion protein comprising beta-secretase fused
CC to a heterologous amino acid sequence, and a method for modulating the
CC levels of beta-secretase polypeptide in a mammal comprising administering
CC the polynucleotide sequence. Beta-secretase exhibits neuroprotective and

```


CC nootropic activity. The beta-secretase nucleotide sequence may be used to
CC map locations of the beta-secretase gene and related genes on chromosomes
CC and as hybridization probes in diagnostic assays to test for the presence
CC of beta-secretase DNA or RNA, such as in Alzheimer's disease, Down's
CC syndrome, and amyloid angiopathy. The nucleotide sequence may also be
CC used as anti-sense inhibitors of beta-secretase expression, in gene
CC therapy of Alzheimer's disease, and for the identification of compounds
CC that modulate beta-secretase activity. Antibodies to the beta-secretase
CC protein may be used for in vitro and in vivo diagnostic purposes to
CC detect the presence of beta-secretase polypeptide in a body fluid or cell
CC sample. The present sequence represents a beta-secretase substrate
CC peptide.
CC
XX
SQ Sequence 8 AA;
Query Match 100.0%; Score 20; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VKMD 4
| | | |
Db 2 VKMD 5
RESULT 9
AAE10660
ID AAE10660 standard; peptide: 8 AA.
XX
AC AAE10660;
XX
DT 10-DEC-2001 (first entry)
XX
DE Human Aspartyl protease-1 (hu-Asp-1) beta-secretase, wild-type peptide.
XX
KW Human; aspartyl protease 1; Aspl; amyloid precursor protein; APP;
KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
KW amyloid plaque; neuronal loss; proteolytic; nootropic; neuroprotective.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Cleavage-site 4..5
XX
PN GB357767-A.
PD 04-JUL-2001.
XX
PE 22-SEP-2000; 2000GB-0023315.
XX
PR 23-SEP-1999; 99US-0155493.
PR 23-SEP-1999; 99US-0404133.
PR 23-SEP-1999; 99MO-US20881.
PR 13-OCT-1999; 99US-0416901.
PR 06-DEC-1999; 99US-0169232.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Bienkowski MJ, Gurney M;
XX
DR WPI: 2001-444208/48.
XX
PT Polypeptide comprising fragments of human aspartyl protease with
PT amyloid precursor protein processing activity and alpha-secretase
PT activity, for identifying modulators useful in treating Alzheimer's
PT disease -
XX
PS Example 15; Page 92; 187pp; English.
XX
CC The patent discloses human aspartyl protease 1 (hu-Asp1) or modified
CC Asp1 proteins which lack transmembrane domain or amino terminal
CC domain or cytoplasmic domain and retains alpha-secretase activity
CC and amyloid protein precursor (APP) processing activity. The proteins
CC of the invention are useful for assaying hu-Asp1 alpha-secretase

CC activity, which in turn is useful for identifying modulators of
CC hu-Asp1 alpha-secretase activity, where modulators that increase
CC hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's
CC disease (AD) which causes progressive dementia with consequent
CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
CC neuronal loss. Hu-Asp1 protease substrate is useful for assaying
CC hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with
CC the substrate under acidic conditions and determining the level of
CC hu-Asp1 proteolytic activity. The present sequence is human aspartyl
CC protease-1 (hu-Asp-1) beta-secretase, wild-type peptide which is used
CC for determining the enzymatic activity of Asp1 protein lacking a
CC transmembrane (TM) domain and containing (His)6 tag.
CC
XX
SQ Sequence 8 AA;
Query Match 100.0%; Score 20; DB 22; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VKMD 4
| | | |
Db 2 VKMD 5
RESULT 10
AAE06902
ID AAE06902 standard; peptide: 8 AA.
XX
AC AAE06902;
XX
DT 23-OCT-2001 (first entry)
XX
DE Human amyloid precursor protein (APP) substrate peptide.
XX
KW Human; aspartyl protease 2; Asp 2; beta-amyloid precursor protein; APP;
KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; nootropic;
KW neuroprotective; antisense therapy; gene therapy.
OS Homo sapiens.
XX
FH WO200150829-A2.
XX
PN 19-JUL-2001.
PD 09-MAY-2001; 2001WO-IB00799.
XX
PE 09-MAY-2001; 2001WO-IB00799.
XX
PR 09-MAY-2001; 2001WO-IB00799.
XX
PA (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX
PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;
XX
DR WPI: 2001-483072/52.
XX
PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity -
XX
PS Claim 128; Page 101; 185pp; English.
XX
CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
CC precursor protein (APP) isoforms and their corresponding DNA molecules.
CC Human aspartyl proteases can act as beta-secretase proteases useful for
CC treating Alzheimer's disease. APP isoforms are useful for identifying
CC modulators of amyloid-beta peptide production, for use in designing
CC therapeutics for the treatment and prevention of Alzheimer's disease,
CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis

CC and neuronal loss. APP isoforms are also used in methods for identifying
 CC inhibitors and modulators of human Asp2 activity. The invention relates
 CC to a method for identifying agents that modulate the activity of human
 CC aspartyl1 protease Asp2. Amyloid-beta peptides obtained from APP are used
 CC as a means to screen in cellular assays for the inhibitors of beta- and
 CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
 CC polymerase chain reactions (PCR). The probes are useful for detecting
 CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
 CC blots. The present sequence is human amyloid precursor protein (APP)
 CC substrate peptide related to the invention.

XX Sequence 8 AA:

Query Match Best Local Similarity 100.0%; Score 20; DB 22; Length 8;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
 ||||
 Db 2 VKMD 5

RESULT 11
 AAU06631
 ID AAU06631 standard; Peptide: 8 AA.

XX AC AAU06631;

XX DT 24-OCT-2001 (first entry)

XX DE Beta secretase substrate peptide.

XX KM Aspartyl1 protease; Asp2; beta-secretase; nototropic;

KM neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
 KM amyloid-beta; Abeta; Beta secretase substrate peptide.

XX OS Synthetic.

XX PN WO200149098-A2.

XX PD 12-JUL-2001.

XX PF 09-MAY-2001; 2001WO-IB00798.

XX PR 09-MAY-2001; 2001WO-IB00798.

XX PA (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

PA (YANR/) YAN R.

PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;

XX WPI; 2001-502549/55.

XX DR Novel purified polypeptide comprising fragment of mammalian aspartyl1
 XX PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 XX PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 XX PT activity

XX Claim 88; Page 94; 185pp; English.

XX The invention relates to a purified polypeptide comprising a fragment of
 CC mammalian aspartyl1 protease (Asp2) protein which lacks the Asp2
 CC transmembrane domain and the Asp2 protein, and where the polypeptide and
 CC the fragment retain the beta-secretase activity of the mammalian Asp2
 CC protein. The invention also details polynucleotides for the APP
 CC proteins and vectors expressing them, and a polypeptide (isoform of
 CC amyloid protein precursor (APP) comprising the amino acid sequence of an
 CC APP or its fragment containing an APP cleavage site recognizable by a
 CC mammalian beta-secretase, and further comprising two lysine residues at
 CC the carboxyl terminus of the amino acid sequence of the mammalian APP or

CC APP fragment. Also included in the invention are methods of identifying
 CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
 CC useful for treating Alzheimer's disease. APP is useful in methods for
 CC identifying inhibitors or modulators of human Asp2 activity and
 CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
 CC therapeutics for the treatment or prevention of Alzheimer's disease.
 CC APP comprising the APP-SW-beta-secretase peptide sequence (NLDA), which
 CC is associated with increased levels of Abeta processing is useful in
 CC assays relating the Alzheimer's research. The expression vector is useful
 CC for recombinantly expressing APP. Nucleic acids that hybridise to
 CC APP oligonucleotides are useful as probes or primers. The probes are
 CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
 CC Northern and Southern blots. The present sequence is a beta secretase
 CC substrate peptide.

XX Sequence 8 AA:

Query Match Best Local Similarity 100.0%; Score 20; DB 22; Length 8;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
 ||||
 Db 2 VKMD 5

RESULT 12
 AAU06635
 ID AAU06635 standard; Peptide: 8 AA.

XX AC AAU06635;

XX DT 24-OCT-2001 (first entry)

XX DE Synthetic fluorescent Asp2 substrate.

XX KM Aspartyl1 protease; Asp2; beta-secretase; nototropic;

KM neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
 KM amyloid-beta; Abeta.

XX OS Synthetic.

XX PN WO200149098-A2.

XX PD 12-JUL-2001.

XX PF 09-MAY-2001; 2001WO-IB00798.

XX PR 09-MAY-2001; 2001WO-IB00798.

XX PA (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

PA (YANR/) YAN R.

PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;

XX WPI; 2001-502549/55.

XX Novel purified polypeptide comprising fragment of mammalian aspartyl1
 XX PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 XX PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 XX PT activity

PS Example 12; Page 81; 185pp; English.
XX
CC The invention relates to a purified polypeptide comprising a fragment of
CC mammalian aspartyl protease (Asp2) protein which lacks the Asp2
CC transmembrane domain and the Asp2 protein, and where the polypeptide and
CC the fragment retain the beta-secretase activity of the mammalian Asp2
CC protein. The invention also details polynucleotides for the Asp
CC proteins and vectors expressing them, and a polypeptide (isoform of
CC amyloid protein precursor (APP)) comprising the amino acid sequence of an
CC APP or its fragment containing an APP cleavage site recognizable by a
CC mammalian beta-secretase, and further comprising two lysine residues at
CC the carboxyl terminus of the amino acid sequence of the mammalian APP or
CC APP fragment. Also included in the invention are methods of identifying
CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
CC useful for treating Alzheimer's disease. APP is useful in methods for
CC identifying inhibitors or modulators of human Asp2 activity and
CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
CC therapeutics for the treatment or prevention of Alzheimer's disease.
CC APP comprising the APP-Su-beta-secretase peptide sequence (NDA), which
CC is associated with increased levels of Abeta processing is useful in
CC assays relating to the Alzheimer's research. The expression vector is useful
CC for recombinantly expressing APP. Nucleic acids that hybridize to
CC APP oligonucleotides are useful as probes or primers. The probes are
CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC Northern and Southern blots. The present sequence is a synthetic
CC fluorescent substrate used to assay Asp2.
XX
SQ Sequence 8 AA:
Query Match 100.0%; Score 20; DB 22; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VKMD 4
 ||||
 2 VKMD 5
Db
RESULT 13
AAU07230
ID AAU07230 standard; Peptide: 8 AA.
XX
AC AAU07230;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human beta-amyloid protein precursor; APP-beta secretase site peptide #3.
XX
KW Human; aspartyl protease 1; Asp-1; neurotrophic; neuroprotective;
KW aspartyl protease 2; Asp2; amyloid protein precursor; APP;
KW beta-secretase; Alzheimer's disease; App-beta.
XX
OS Homo sapiens.
XX
FN WO200149097-A2.
XX
PD 12-JUL-2001.
XX
PF 09-MAY-2001; 2001WO-IB00797.
XX
PR 09-MAY-2001; 2001WO-IB00797.
XX
PA (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX
PI Blenkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;
XX WPI: 2001-502548/55.
XX
PT Novel purified polypeptide comprising fragment of mammalian aspartyl

PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity
XX
XX Claim 88; Page 94; 185pp; English.
XX
CC The invention relates to a novel purified polypeptide comprising a
CC fragment of mammalian aspartyl protease 2 (Asp2) protein which lacks the
CC Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide
CC and the fragment retain the beta-secretase activity of the mammalian Asp2
CC protein. Also included is an isoform of amyloid protein precursor (APP)
CC comprising the amino acid sequence of a APP or its fragment containing
CC an APP cleavage site recognizable by a mammalian beta-secretase, and
CC further comprising two lysine residues at the carboxyl terminus of the
CC amino acid sequence of the mammalian APP or APP fragment. The
CC polypeptides are used for assaying for modulators of beta-secretase
CC activity; identifying agents that inhibit the APP processing activity
CC of human Asp2 aspartyl protease (Hu-Asp2); identifying agents that
CC modulate the activity of Asp2; and for reducing cellular production of
CC amyloid beta (Abeta) from APP. Agents identified by the above methods
CC are useful for treating Alzheimer's disease; and for identifying
CC modulators of amyloid-beta (Abeta) peptide production, for use in
CC designing therapeutics for the treatment or prevention of Alzheimer's
CC disease. Probes and primers derived from APP nucleic acid sequences
CC are useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC Northern and Southern blots. The present sequence represents the
CC amino acid sequence of human amyloid protein precursor, APP-beta
CC secretase site peptide substrate #3 used in assays of human Asp2 beta-
CC secretase activity.
XX
SQ Sequence 8 AA:
Query Match 100.0%; Score 20; DB 22; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VKMD 4
 ||||
 2 VKMD 5
Db
RESULT 14
AAE02612
ID AAE02612 standard; peptide: 8 AA.
XX
AC AAE02612;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human Aspartyl protease-1 (Hu-Asp-1) beta-secretase, wild-type peptide.
XX
KW Human; alpha-secretase; amyloid precursor protein; APP; therapy;
KW Alzheimer's disease; antialzheimer's; aspartyl protease 1; Asp1;
KW beta-secretase.
XX
OS Homo sapiens.
XX
FN WO200123533-A2.
XX
PD 05-APR-2001.
XX
PF 22-SEP-2000; 2000WO-US26080.
XX
PR 23-SEP-1999; 99US-0155493.
PR 23-SEP-1999; 99WO-US20881.
PR 13-OCT-1999; 99US-0416901.
PR 06-DEC-1999; 99US-0169232.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.

PI Gurney M, Blenkowski MJ;
 XX
 DR WPI; 2001-290516/30.
 XX
 PT Enzymes that cleave the alpha-secretase site of the amyloid precursor
 XX protein, useful for the treatment of Alzheimer's disease -
 XX
 XX Example 15; Page 94; 189pp; English.
 CC The present invention relates to enzymes for cleaving the alpha-
 CC secretase site of the amyloid precursor protein (APP) and methods of
 CC identifying those enzymes. The methods may be used to identify enzymes
 CC that may be used to cleave the alpha-secretase cleavage site of the APP
 CC protein. The enzymes may be used to treat or modulate the progress of
 CC Alzheimer's disease. The present sequence is human Aspartyl protease-1
 CC (hu-Asp-1) beta-secretase, wild-type peptide which is used for
 CC determining the enzymatic activity of Asp-1 deltatm (His)6 protein.
 CC
 XX Sequence 8 AA:
 SQ
 Query Match 100.0%; Score 20; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VKMD 4
 IIII
 DB 2 VKMD 5
 DB
 RESULT 15
 AAW82084
 ID AAW82084 standard; peptide; 9 AA.
 XX
 AC AAW82084;
 XX
 DT 18-FEB-1999 (first entry)
 XX
 DE Fluorogenic protease indicator protease binding peptide #62.
 XX
 KW Protease activity; fluorophore; detection; fluorogenic; cellular uptake;
 KW conformation change.
 XX
 OS Synthetic.
 OS
 PN WO9837226-A1.
 XX
 PD 27-AUG-1998.
 XX
 PF 20-FEB-1998; 98WO-US03000.
 XX
 PR 20-FEB-1997; 97US-0802981.
 XX
 PA (ONCO-) ONCOIMMUNIN INC.
 XX
 PI Komoriya A, Packard BS;
 XX
 DR WPI; 1998-467579/40.
 XX
 PT New fluorogenic compositions - containing 2 fluorophores separated
 PT by a peptide comprising a protease binding site, used for detecting
 PT protease activity in samples.
 XX
 PS Claim 4; Page 77; 90pp; English.
 XX
 CC AAW82023-W82240 are peptides used in the construction of a fluorogenic
 CC composition which is used for the detection of protease activity in
 CC biological samples. The products can be used for the detection of
 CC conformation changes in nucleic acids, oligosaccharides,
 CC polysaccharides, proteins, lipids, phospholipids, glycolipids,
 CC glycoproteins, steroids or polymers. In addition, attachment of a
 CC hydrophobic group to a molecule can be used to enhance uptake by cells.
 CC The composition is composed of P = peptide comprising a protease binding
 CC site for the protease, F1, F2 peptides = fluorophores where F1 is

CC attached to the amino terminal amino acid and F2 is attached to the
 CC carboxyl terminal amino acid and S1, S2 peptides = when present, are
 CC peptide spacers where S1, when present, is attached to the amino terminal
 CC acid, and S2, when present, is attached to the carboxyl terminal amino
 CC acid.
 XX
 SQ Sequence 9 AA:
 Query Match 100.0%; Score 20; DB 19; Length 9;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VKMD 4
 IIII
 DB 3 VKMD 6
 DB
 Search completed: October 30, 2002, 12:27:10
 Job time : 2.60197 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:24:20 ; Search time 0.570025 Seconds
(without alignments)
171.400 Million cell updates/sec

Title: US-09-724-571-103

Perfect score: 20

Sequence: 1 VKMD 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	US-09-173-887-1	Sequence 1, Appli
2	20	100.0	4	US-09-294-987-2	Sequence 2, Appli
3	20	100.0	4	US-09-054-334-5	Sequence 5, Appli
4	20	100.0	5	US-08-480-498-1	Sequence 1, Appli
5	20	100.0	5	US-08-659-984A-13	Sequence 13, Appli
6	20	100.0	5	US-08-660-531-13	Sequence 13, Appli
7	20	100.0	5	US-09-054-334-1	Sequence 1, Appli
8	20	100.0	7	US-08-792-553-10	Sequence 10, Appli
9	20	100.0	8	US-07-965-971-4	Sequence 4, Appli
10	20	100.0	8	US-08-143-697-4	Sequence 4, Appli
11	20	100.0	8	US-08-440-261-4	Sequence 4, Appli
12	20	100.0	8	US-08-440-423-4	Sequence 4, Appli
13	20	100.0	8	US-08-846-444-4	Sequence 4, Appli
14	20	100.0	8	PCT-US94-07043A-4	Sequence 4, Appli
15	20	100.0	9	US-08-802-981-221	Sequence 221, App
16	20	100.0	9	US-08-802-981-222	Sequence 222, App
17	20	100.0	9	US-08-802-981-222	Sequence 222, App
18	20	100.0	10	US-09-294-987-6	Sequence 7, Appli
19	20	100.0	11	US-08-025-321C-1	Sequence 1, Appli
20	20	100.0	12	PCT-US94-07043A-7	Sequence 2, Appli
21	20	100.0	16	PCT-US94-07043A-2	Sequence 3, Appli
22	20	100.0	18	US-08-149-975A-3	Sequence 114, App
23	20	100.0	21	US-08-802-981-114	Sequence 115, App
24	20	100.0	21	US-08-802-981-115	Sequence 11, Appli
25	20	100.0	27	US-08-141-324-11	Sequence 11, Appli
26	20	100.0	27	US-08-541-902-11	Sequence 35, App
27	20	100.0	27	US-08-936-165A-335	

28	20	100.0	45	1	US-08-462-859A-5	Sequence 5, Appli
29	20	100.0	45	1	US-08-123-659A-5	Sequence 5, Appli
30	20	100.0	45	1	US-08-464-247A-5	Sequence 5, Appli
31	20	100.0	45	1	US-08-464-248A-5	Sequence 5, Appli
32	20	100.0	53	4	US-09-173-887-5	Sequence 5, Appli
33	20	100.0	53	4	US-09-294-987-1	Sequence 1, Appli
34	20	100.0	58	1	US-08-371-930-25	Sequence 25, Appli
35	20	100.0	58	5	PCT-US94-01712-25	Sequence 25, Appli
36	20	100.0	59	1	US-08-484-969-3	Sequence 3, Appli
37	20	100.0	59	1	US-08-472-627-3	Sequence 3, Appli
38	20	100.0	59	1	US-08-388-463-3	Sequence 3, Appli
39	20	100.0	63	1	US-08-462-859A-3	Sequence 3, Appli
40	20	100.0	63	1	US-08-462-859A-4	Sequence 4, Appli
41	20	100.0	63	1	US-08-123-659A-3	Sequence 4, Appli
42	20	100.0	63	1	US-08-123-659A-4	Sequence 4, Appli
43	20	100.0	63	1	US-08-464-247A-3	Sequence 4, Appli
44	20	100.0	63	1	US-08-464-247A-4	Sequence 4, Appli
45	20	100.0	63	1	US-08-464-248A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-173-887-1
; Sequence 1, Application US/09173887
; Patent No. 6245884
; GENERAL INFORMATION:
; APPLICANT: Hook, Vivian Y.H.
; TITLE OF INVENTION: SECRETASES RELATED TO ALZHEIMER'S DEMENTIA
; FILE REFERENCE: P-AS 3337
; CURRENT APPLICATION NUMBER: US/09/173,887
; CURRENT FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: mammalian
US-09-173-887-1

Query Match 100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
DB 1 VKMD 4

RESULT 2
US-09-294-987-2
; Sequence 2, Application US/09294987
; Patent No. 6313268
; GENERAL INFORMATION:
; APPLICANT: Hook, Vivian Y.H.
; TITLE OF INVENTION: SECRETASES RELATED TO ALZHEIMER'S DEMENTIA
; FILE REFERENCE: P-AS 3515
; CURRENT APPLICATION NUMBER: US/09/294,987
; CURRENT FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: US 09/173,887
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 4
; TYPE: PRT
; ORGANISM: mammalian
US-09-294-987-2

Query Match 100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
Db 1 VKMD 4

RESULT 3
US-09-054-334-5
; Sequence 5, Application US/09054334
; Patent No. 6329163
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Jacobson-Croak, Kirsten L.
; APPLICANT: Sinha, Sukanto
; TITLE OF INVENTION: Assays for Detecting Beta-Secretase
; TITLE OF INVENTION: Inhibition
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/054,334
; FILING DATE: 02-APR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,152
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 015270-00282005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-054-334-5

Query Match 100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
Db 1 VKMD 4

RESULT 4
US-08-480-498-1
; Sequence 1, Application US/08480498
; Patent No. 5744346
; GENERAL INFORMATION:
; APPLICANT: Chrysler, Susanna M.S.
; APPLICANT: Sinha, Sukanto
; APPLICANT: Keim, Pamela S.
; APPLICANT: Anderson, John P.
; TITLE OF INVENTION: Beta-Secretase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,498
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 015270-002200
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-480-498-1

Query Match 100.0%; Score 20; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
Db 1 VKMD 4

RESULT 5
US-08-659-984A-13
; Sequence 13, Application US/08659984A
; Patent No. 5942400
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Sinha, Sukanto
; APPLICANT: Jacobson-Croak, Kirsten L.
; TITLE OF INVENTION: Assays for Detecting Beta-Secretase
; TITLE OF INVENTION: Inhibition
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Ctr., 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,984A
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,152
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541

REFERENCE/DOCKET NUMBER: 15270-002810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-659-984A-13

Query Match 100.0%; Score 20; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKMD 4
Db 1 VKMD 4

RESULT 6
US-08-660-531-13
Sequence 13, Application US/08660531
Patent No. 6221645
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Keim, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,531
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,498
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-660-531-13

Query Match 100.0%; Score 20; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKMD 4

Db 1 VKMD 4

RESULT 7
US-09-054-334-1
Sequence 1, Application US/09054334
Patent No. 6329163
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Jacobson-Croak, Kirsten L.
APPLICANT: Sinha, Sukanto
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
TITLE OF INVENTION: Inhibition
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,334
FILING DATE: 02-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 015270-002820US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-054-334-1

Query Match 100.0%; Score 20; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKMD 4
Db 1 VKMD 4

RESULT 8
US-08-792-553-10
Sequence 10, Application US/08792553
Patent No. 5981200
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Helm, Roger
TITLE OF INVENTION: Tandem Fluorescent Protein Constructs
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & RICHARDSON P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California

COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/792,553
APPLICATION NUMBER: US/08/792,553
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lisa A. Halle, Ph.D.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/041001/UC 96-160-2
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-792-553-10

Query Match 100.0%; Score 20; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
Db 1 VKMD 4

RESULT 9
US-07-965-971-4
Sequence 4, Application US/07965971
Patent No. 5441870
GENERAL INFORMATION:
APPLICANT: Seubert, Peter A.
APPLICANT: Schenk, Dale B.
APPLICANT: Fritz, Lawrence C.
TITLE OF INVENTION: Methods and Compositions for Monitoring
TITLE OF INVENTION: Cellular Processing of Beta-Amyloid Precursor Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/965,971
FILING DATE: 26-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/868,949
FILING DATE: 15-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-4-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-965-971-4

Query Match 100.0%; Score 20; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
Db 5 VKMD 8

RESULT 10
US-08-143-697-4
Sequence 4, Application US/08143697
Patent No. 5604102
GENERAL INFORMATION:
APPLICANT: Seubert, Peter A.
APPLICANT: Schenk, Dale B.
APPLICANT: Fritz, Lawrence C.
TITLE OF INVENTION: Methods of Screening for Beta-Amyloid
TITLE OF INVENTION: Peptide Production Inhibitors
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/143,697
FILING DATE: 27-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,971
FILING DATE: 26-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/868,949
FILING DATE: 15-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-4-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-143-697-4

Query Match 100.0%; Score 20; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
Db 1 VKMD 4

Db 5 VKMD 8

RESULT 11
US-08-440-261-4
; Sequence 4, Application US/08440261
; Patent No. 5605811
; GENERAL INFORMATION:
; APPLICANT: Seubert, Peter A.
; APPLICANT: Schenk, Dale B.
; APPLICANT: Filtz, Lawrence C.
; TITLE OF INVENTION: Methods and Compositions for Monitoring
; TITLE OF INVENTION: Cellular Processing of Beta-Amyloid Precursor Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,261
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,971
; FILING DATE: 26-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-4-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-440-261-4

Query Match 100.0%; Score 20; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
Db 5 VKMD 8

RESULT 12
US-08-440-423-4
; Sequence 4, Application US/08440423
; Patent No. 5721130
; GENERAL INFORMATION:
; APPLICANT: Seubert, Peter A.
; APPLICANT: Schenk, Dale B.
; APPLICANT: Filtz, Lawrence C.
; TITLE OF INVENTION: Methods and Compositions for Monitoring
; TITLE OF INVENTION: Cellular Processing of Beta-Amyloid Precursor Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco

STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,423
FILING DATE: 12-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,971
FILING DATE: 26-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-4-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-440-423-4

Query Match 100.0%; Score 20; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
Db 5 VKMD 8

RESULT 13
US-08-846-444-4
; Sequence 4, Application US/08846444
; Patent No. 6018024
; GENERAL INFORMATION:
; APPLICANT: Seubert, Peter A.
; APPLICANT: Schenk, Dale B.
; APPLICANT: Filtz, Lawrence C.
; TITLE OF INVENTION: Methods and Compositions for Monitoring
; TITLE OF INVENTION: Cellular Processing of Beta-Amyloid Precursor Protein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/846,444
; FILING DATE: No. 6018024 yet assigned
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/868,949
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,971
; FILING DATE: 26-OCT-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US93/01817
FILING DATE: 03-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/440,423
FILING DATE: 12-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 015270-00045005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-846-444-4

Query Match
Best Local Similarity 100.0%; Score 20; DB 3; Length 8;
Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
1111
DB 5 VKMD 8

RESULT 14
PCT-US94-07043A-4
SEQUENCE 4, Application PC/TUS9407043A
GENERAL INFORMATION:
APPLICANT: Tamburini, Paul P.; Benz, G nter; H blich,
APPLICANT: Dieter, Robert N.; Koenig, Gerhard
TITLE OF INVENTION: CATHEPSIN D IS AN AMYLOIDGENIC
TITLE OF INVENTION: PROTEASE IN ALZHEIMER S DISEASE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Miles Inc.
STREET: 400 Morgan Lane
CITY: West Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06516
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage
COMPUTER: Sharp PC 4600
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07043A
FILING DATE: June 21, 1994
CLASSIFICATION:
PRIOR APPLICATION DATA: PCT/US93/10889
APPLICATION NUMBER: PCT/US93/10889
FILING DATE: November 12, 1993
PRIOR APPLICATION DATA: 07/995,660
APPLICATION NUMBER: 07/995,660
FILING DATE: December 16, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/880,914
FILING DATE: May 11, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Pamela A. Simonton
REGISTRATION NUMBER: 31,060
REFERENCE/DOCKET NUMBER: MTI 224.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 937-2340
TELEFAX: (203) 937-2795
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US94-07043A-4

Query Match
Best Local Similarity 100.0%; Score 20; DB 5; Length 8;
Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
1111
DB 5 VKMD 8

RESULT 15
US-08-802-981-221
SEQUENCE 221, Application US/08802981
PATENT NO. 6037137
GENERAL INFORMATION:
APPLICANT: Komoriya, Akira
APPLICANT: Packard, Beverly S.
TITLE OF INVENTION: Compositions for the Detection of Enzyme
TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/802,981
FILING DATE: 20-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 016865-00030005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 221:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-802-981-221

Query Match
Best Local Similarity 100.0%; Score 20; DB 3; Length 9;
Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
1111
DB 3 VKMD 6

Search completed: October 30, 2002, 12:32:36
Job time: 1.57002 secs

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2002, 09:13:36 ; Search time 50.29 Seconds

(without alignments)
1568.616 Million cell updates/sec

Title: US-09-724-571-43

Perfect score: 2419
Sequence: 1 ETDEPEPEPRGRSGFEVEMVD.....CLRCRCQHQHDDFADISILK 456Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2419	100.0	532	4 Q9ULS1	Q9ULS1 homo sapien
2	2264.5	93.6	476	4 Q9BYC1	Q9BYC1 homo sapien
3	2161	89.3	457	4 Q9BYC0	Q9BYC0 homo sapien
4	2016.5	83.4	432	4 Q9BYB9	Q9BYB9 homo sapien
5	1407	58.2	266	11 Q9UCU5	Q9UCU5 mus musculu
6	1160.5	48.0	439	4 Q9H2V8	Q9H2V8 homo sapien
7	1146.5	47.4	514	11 Q9JL18	Q9JL18 mus musculu
8	972.5	40.2	468	4 Q9NZL2	Q9NZL2 homo sapien
9	967.5	40.0	396	4 Q9NZL1	Q9NZL1 homo sapien
10	712.5	29.5	213	4 Q9P0D2	Q9P0D2 homo sapien
11	596.5	24.7	255	11 Q9R1P7	Q9R1P7 mus musculu
12	332.5	13.7	391	5 Q9VKP6	Q9VKP6 drosophila
13	332	13.7	354	5 Q9GXY7	Q9GXY7 boophilus m
14	312.5	12.9	386	6 Q9BUC5	Q9BUC5 bos taurus
15	308	12.7	384	13 Q9DEC2	Q9DEC2 xenopus lae
16	308	12.7	385	13 Q9DEC4	Q9DEC4 rana catesb

17	305	12.6	386	6 Q9GMV7	Q9GMV7 rhinolophus
18	305	12.6	387	6 Q9GMV8	Q9GMV8 scroex ungui
19	304.5	12.6	372	5 Q9YLK3	Q9YLK3 drosophila
20	304.5	12.6	387	13 Q9DDV5	Q9DDV5 salvellinus
21	304	12.6	387	6 Q9GMV9	Q9GMV9 suncus murti
22	303.5	12.5	383	13 Q9DE45	Q9DE45 salvelinus
23	303	12.5	383	13 Q9DEC3	Q9DEC3 xenopus lae
24	302.5	12.5	376	13 Q9PUR8	Q9PUR8 pseudopleur
25	301	12.4	382	13 Q9PRG9	Q9PRG9 gallus galli
26	301	12.4	423	5 Q9YKP7	Q9YKP7 drosophila
27	298.5	12.3	384	13 Q913Z2	Q913Z2 rana catesb
28	295.5	12.2	386	6 Q9GMV6	Q9GMV6 canis famill
29	294	12.2	396	13 Q934Z8	Q934Z8 chionodraco
30	290.5	12.0	381	6 Q9GK11	Q9GK11 camelus dro
31	288	11.9	399	13 Q93458	Q93458 podarcis sl
32	287.5	11.9	444	5 Q21966	Q21966 caenorhabdi
33	284	11.7	398	13 P87370	P87370 oncorhynch
34	284	11.7	427	5 P91802	P91802 schistosoma
35	281	11.6	378	13 Q9PUR9	Q9PUR9 pseudopleur
36	279.5	11.6	390	6 Q9GK10	Q9GK10 camelus dro
37	278	11.5	370	6 Q9TWT1	Q9TWT1 bos taurus
38	278	11.5	399	13 Q9DD89	Q9DD89 brachydanio
39	278	11.5	422	5 Q96906	Q96906 onchocerca
40	277.5	11.5	380	6 Q28950	Q28950 sus scrofa
41	277	11.5	446	5 Q9N9H3	Q9N9H3 necator ame
42	276	11.4	389	13 Q9PMK1	Q9PMK1 gallus galli
43	276	11.4	389	13 Q9W643	Q9W643 gallus galli
44	276	11.4	396	13 Q9DEX3	Q9DEX3 clupea hare
45	275	11.4	385	6 Q29080	Q29080 sus scrofa

ALIGNMENTS

RESULT 1
ID Q9ULS1 PRELIMINARY: PRT: 532 AA.
AC Q9ULS1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE KIAA1149 PROTEIN (FRAGMENT).
GN KIAA1149.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=20039618; PubMed=10574461;
RA Hiroseawa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;
RT "Characterization of cDNA clones selected by the Genemark analysis
from size-fractionated cDNA libraries from human brain.";
RL DNA Res. 6:329-336(1999).
DR EMBL: AB032975; BAA6463.2; -
DR HSSP: P56272; IAW5.
DR MEROPS: A01.004; -
DR InterPro: IPR001969; Asp-protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF000026; asp. 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 532 AA; 58720 MW; 96B135DD5FED2E8 CRC64;

Query Match 100.0%; Score 2419; DB 4; Length 532;
Best Local Similarity 100.0%; Pred. No. 7 3e-198;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ETDEPEPEPRGRSGFEVEMVDNLKRGSGGYVEMVTGSPQTLNLTVDGTGSSNFAVGAP 60
|||||

```

Db 77 ETDEPEEPGRGSEFVEMVDNLGRKSGQGYVEMTVGSPQTLNLTVDTGSSNFVAGAAP 136
QY 61 HPELHRYQROLSSYTRDLRKGYYVPTQGWEGELGTDLVSIHPGPNVYRANIAATE 120
Db 137 HPELHRYQROLSSYTRDLRKGYYVPTQGWEGELGTDLVSIHPGPNVYRANIAATE 196
QY 121 SDKPFINSNMEGILGLAYAEIARPDLSLEPFDSLVRKQTHVNFSLQAGAPLNO 180
Db 197 SDKPFINSNMEGILGLAYAEIARPDLSLEPFDSLVRKQTHVNFSLQAGAPLNO 256
QY 181 EYLASVGSMITGGIDHSLYTGSLWYTPIRREMYEVIIVVEINGODLKMDCKEYNDK 240
Db 257 EYLASVGSMITGGIDHSLYTGSLWYTPIRREMYEVIIVVEINGODLKMDCKEYNDK 316
QY 241 STVDSGTINLRPKKVFEEAAYKSIKAASSTKEKPDGFMGLGEOVGMQAGTTPMNIFFPYS 300
Db 317 STVDSGTINLRPKKVFEEAAYKSIKAASSTKEKPDGFMGLGEOVGMQAGTTPMNIFFPYS 376
QY 301 LYLMGEVTVNOSFRITILPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 360
Db 377 LYLMGEVTVNOSFRITILPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 436
QY 361 FDRARRKRIGFSAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDESLMTIAYVMAI 420
Db 437 FDRARRKRIGFSAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDESLMTIAYVMAI 496
QY 421 CALFMLPCLIMVCOMRCLRCLROQHDHDFADISILK 456
Db 497 CALFMLPCLIMVCOMRCLRCLROQHDHDFADISILK 532

RESULT 2
QYBC1 PRELIMINARY: PRT: 476 AA.
AC Q9BYC1.
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE BETA-SITE APP CLEAVING ENZYME I-476.
CN BACE.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H., Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production.";
RL Neurosci. Lett. 307:9-12(2001).
DR EMBL; AB050436; BAB40931.1; -.
DR HSSP; P32329; 1YPS.
DR InterPro; IPR001969; Asp--protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 476 AA; 52907 MW; 6C8C87F8A953AF66 CRC64;

Query Match 93.6%; Score 2264.5; DB 4; Length 476;
Best Local Similarity 94.5%; Pred. No. 9.5e-185;
Matches 431; Conservative 0; Mismatches 0; Indels 25; Gaps 1;

```

```

Db 106 HPELHRYQROLSSYTRDLRKGYYVPTQGWEGELGTDLVSIHPGPNVYRANIAATE 165
QY 121 SDKPFINSNMEGILGLAYAEIARPDLSLEPFDSLVRKQTHVNFSLQAGAPLNO 180
Db 166 SDKPFINSNMEGILGLAYAEIARPDLSLEPFDSLVRKQTHVNFSLQAGAPLNO 200
QY 181 EYLASVGSMITGGIDHSLYTGSLWYTPIRREMYEVIIVVEINGODLKMDCKEYNDK 240
Db 201 EYLASVGSMITGGIDHSLYTGSLWYTPIRREMYEVIIVVEINGODLKMDCKEYNDK 260
QY 241 STVDSGTINLRPKKVFEEAAYKSIKAASSTKEKPDGFMGLGEOVGMQAGTTPMNIFFPYS 300
Db 261 STVDSGTINLRPKKVFEEAAYKSIKAASSTKEKPDGFMGLGEOVGMQAGTTPMNIFFPYS 320
QY 301 LYLMGEVTVNOSFRITILPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 360
Db 321 LYLMGEVTVNOSFRITILPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 380
QY 361 FDRARRKRIGFSAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDESLMTIAYVMAI 420
Db 381 FDRARRKRIGFSAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDESLMTIAYVMAI 440
QY 421 CALFMLPCLIMVCOMRCLRCLROQHDHDFADISILK 456
Db 441 CALFMLPCLIMVCOMRCLRCLROQHDHDFADISILK 476

RESULT 3
QYBC0 PRELIMINARY: PRT: 457 AA.
AC Q9BYC0.
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE BETA-SITE APP CLEAVING ENZYME I-457 (BETA-SITE APP CLEAVING ENZYME
DE TYPE C).
CN BACE.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H., Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production.";
RL Neurosci. Lett. 307:9-12(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=EXOCRINE PANCREAS;
RA Zaccchetti D., De Pietri Tonelli D., Schnurbus R.;
RT "New beta-site APP cleaving enzyme isoform (BACE-1c) obtained from
RT human pancreas.";
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB050437; BAB40932.1; -.
DR EMBL; AF338817; AAK38375.1; -.
DR HSSP; P32329; 1YPS.
DR InterPro; IPR001969; Asp--protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 457 AA; 51068 MW; C794C9A9E8B5FE7A2 CRC64;

Query Match 89.3%; Score 2161; DB 4; Length 457;
Best Local Similarity 90.4%; Pred. No. 6.1e-176;
Matches 412; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

```

```
|||||
Db 46 ETDEPEEPERRRSSFEVMVNLGRKSGGGYVEMTVGSPQTLINIIIVDGSSNFAGVAAAP 105
QY 61 HPELHRYRQRLSSYTDLRKGYVVPYTGKMEGELGTDLVSPHGPNTVVRANIAATE 120
Db 106 HPELHRYRQRLSSYTDLRKGYVVPYTGKMEGELGTDLVSPHGPNTVVRANIAATE 145
QY 121 SDFEFLNGSMWEGILGLAVAEIARPDSDLEPFDSLVKQTHVNLFSLOLCGAGPFLNS 180
Db 146 -----PDOSLEPFDSLVKQTHVNLFSLOLCGAGPFLNS 181
QY 181 EVLASVGGSMIIIGIDHSLYTSLWYTPIRREMYEVIIVREINQDLKMDCKEYNDK 240
Db 182 EVLASVGGSMIIIGIDHSLYTSLWYTPIRREMYEVIIVREINQDLKMDCKEYNDK 241
QY 241 SIYDGGTTLRLPKKVFEEAAVKSIAASTEKFPDGFMLGEOLVQWQAGTTPMNIFFVIS 300
Db 242 SIYDGGTTLRLPKKVFEEAAVKSIAASTEKFPDGFMLGEOLVQWQAGTTPMNIFFVIS 301
QY 301 LYLMGEVYNOSFRITILPQOYLRPVEDVATSODDCYKFAISOSTGTGVAGVIMEGFYV 360
Db 302 LYLMGEVYNOSFRITILPQOYLRPVEDVATSODDCYKFAISOSTGTGVAGVIMEGFYV 361
QY 361 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAI 420
Db 362 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAI 421
QY 421 CALFMLPLCLMWQWRCLRCLRQOHDFADDISLTK 456
Db 422 CALFMLPLCLMWQWRCLRCLRQOHDFADDISLTK 457
```

```
RESULT 4
Q9BYB9 PRELIMINARY; PRT; 432 AA.
ID Q9BYB9;
AC Q9BYB9;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE BETA-SITE APP CLEAVING ENZYME I-432.
GN BACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H., Tadira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production.";
RL Neurosci. Lett. 307:9-12(2001).
DR EMBL; AB050438; BAB40933.1; -.
DR HSSP; P32329; 1YPS.
DR Interpro: IPR001969; Asp_protase.
DR Interpro: IPR001461; Pepsin.
DR Pfam; PF00026; asp.1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN.1.
SO SEQUENCE 432 AA; 48212 MW; 96FC81E6F0EED01B CRC64;
```

```
Query Match 83.4%; Score 2016.5; DB 4; Length 432;
Best Local Similarity 84.9%; Pred. No. 1,2e-163;
Matches 387; Conservative 0; Mismatches 0; Indels 69; Gaps 1;
```

QY 1 ETDEPEEPGRGSGFVEMVNLGRKSGGGYVEMTVGSPQTLINIIIVDGSSNFAGVAAAP 60
|||||
Db 46 ETDEPEEPGRGSGFVEMVNLGRKSGGGYVEMTVGSPQTLINIIIVDGSSNFAGVAAAP 105
|||||
QY 61 HPELHRYRQRLSSYTDLRKGYVVPYTGKMEGELGTDLVSPHGPNTVVRANIAATE 120

```
|||||
Db 106 HPELHRYRQRLSSYTDLRKGYVVPYTGKMEGELGTDLVSPHGPNTVVRANIAATE 145
QY 121 SDFEFLNGSMWEGILGLAVAEIARPDSDLEPFDSLVKQTHVNLFSLOLCGAGPFLNS 180
Db 146 -----LCGAGPFLNS 156
QY 181 EVLASVGGSMIIIGIDHSLYTSLWYTPIRREMYEVIIVREINQDLKMDCKEYNDK 240
Db 157 EVLASVGGSMIIIGIDHSLYTSLWYTPIRREMYEVIIVREINQDLKMDCKEYNDK 216
QY 241 SIYDGGTTLRLPKKVFEEAAVKSIAASTEKFPDGFMLGEOLVQWQAGTTPMNIFFVIS 300
Db 212 SIYDGGTTLRLPKKVFEEAAVKSIAASTEKFPDGFMLGEOLVQWQAGTTPMNIFFVIS 276
QY 301 LYLMGEVYNOSFRITILPQOYLRPVEDVATSODDCYKFAISOSTGTGVAGVIMEGFYV 360
Db 277 LYLMGEVYNOSFRITILPQOYLRPVEDVATSODDCYKFAISOSTGTGVAGVIMEGFYV 336
QY 361 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAI 420
Db 337 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDDESTLMTIAYVMAI 396
QY 421 CALFMLPLCLMWQWRCLRCLRQOHDFADDISLTK 456
Db 397 CALFMLPLCLMWQWRCLRCLRQOHDFADDISLTK 432
```

```
RESULT 5
Q9CUC5 PRELIMINARY; PRT; 266 AA.
ID Q9CUC5;
AC Q9CUC5;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE BETA-SITE APP CLEAVING ENZYME (FRAGMENT).
GN BACE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Mikaldo I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderrelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
RA Wyshnaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK014390; BAB29317.1; -.
DR MEROBS; A01.004; -.
DR MGD; MGI:1346342; Bace.
DR Interpro: IPR001461; Pepsin.
DR Pfam; PF00026; asp.2.
FT NON_TER 1
SO SEQUENCE 266 AA; 30201 MW; B913FDA8ADAB4238 CRC64;
```

```

Query Match          58.2%; Score 1407; DB 11; Length 266;
Best Local Similarity 98.9%; Pred. No. 5.9e-112;
Matches 263; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 191 IIGIDHSLYTGSIMWTPPIRREWEYEVIIIVREINGODLMDCKEYNDKSIYVSGTINL 250
DB 1 IIGIDHSLYTGSIMWTPPIRREWEYEVIIIVREINGODLMDCKEYNDKSIYVSGTINL 60

QY 251 RLPKRVFAAAYKSIKAASTEFKPPGFMGLGBOLYCWOAGTTPMNIPIYISLYLMEVYTNQ 310
DB 61 RLPKRVFAAAYKSIKAASTEFKPPGFMGLGBOLYCWOAGTTPMNIPIYISLYLMEVYTNQ 120

QY 311 SFRITILPQOYLRAPEVDATSDDCYKFAISOSTGTVMGAVIMEGYVVEDRARKRIGF 370
DB 121 SFRITILPQOYLRAPEVDATSDDCYKFAISOSTGTVMGAVIMEGYVVEDRARKRIGF 180

QY 371 AVSACHVHDEFTAAVEGPFYTLDMEDCGYNIPOTDESTIMTIAVMAICALEMPLICL 430
DB 181 AVSACHVHDEFTAAVEGPFYTLDMEDCGYNIPOTDESTIMTIAVMAICALEMPLICL 240

QY 431 MYCOMRCLRCLRQOHDDFADDISLTK 456
DB 241 MYCOMRCLRCLRQOHDDFADDISLTK 266

RESULT 6
Q9H2V8 PRELIMINARY; PRT; 439 AA.
AC Q9H2V8;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE CDAL3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PHOCHROMOCYTOMA;
RA Li Y., Huang Q., Peng, Y., Song H., Yu Y., Xu S., Ren S., Chen Z.,
RA Han Z.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBD databases.
DR EMBL; AF212252; AAG41783.1; -.
DR HSSP; P00797; ZREN.
DR InterPro; IPR001969; Asp_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 3.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 439 AA; 48275 MW; 02EC0E0E50F11602 CRC64;

Query Match          48.0%; Score 1160.5; DB 4; Length 439;
Best Local Similarity 50.1%; Pred. No. 1.3e-90;
Matches 220; Conservative 78; Mismatches 134; Indels 7; Gaps 4;

QY 18 MYDNLGRSGGGYVEMTVGSPPTNLIVDTGSSNFAVGAPHPHLRIYQQLSSTR 77
DB 1 MYDNLGRSGGGYVEMTVGSPPTNLIVDTGSSNFAVGAPHPHLRIYQQLSSTR 60

QY 78 DLKRGVYVPTQKWEGLDGLVSTPHGPNVYVRANIIATSDKFKINGSMWEGILGL 137
DB 61 SKGPDVYVKTQSGWTFGEVEDLVTPKGFNTSFLVNIATIFESNFLLGKIMNGILGL 120

QY 138 AYAEIARPDSDLPEFPDLVKQTHVPMLESLQLCGAGFPLNOSSEVLASVGSMIIGSIDH 197
DB 121 AYATIAFPSSLEFPFSLVTOANIRPVFEMQMGAGLPAVAGS---GTNGSLVLGSIETP 177

QY 198 SLVYTGSLMTPPIRREWEYEVIIIVREINGODLMDCKEYNDKSIYVSGTINLRLPKRV 257
DB 178 SLVYTGSLMTPPIRREWEYEVIIIVREINGODLMDCKEYNDKSIYVSGTINLRLPKRV 237

```

```

QY 258 EAAVKSITKAASSTEFKPPDGMGLBOLYCWOAGTTPMNIPIYISLYLMEVYTNQSFRTIIL 317
DB 238 DAVEAARASLPIEFSDGFWTGSQGLCWNSFPMYFKISITLYLDENSSSFRITIL 297

QY 318 POOYLRAPEVDATSDDCYKFAISOSTGTVMGAVIMEGYVVEDRARKRIGFAVSAACHV 377
DB 298 POLYIQPMGAGLNY-ECYRGISPTNALVIGATVMEGYVIFEDRAQKRVGFAASPCAE 356

QY 378 HDEFTAAVEGPFYTLDMEDCGYNIPOTDESTIMTIAVMAIC-ALFMLPCLMCOWR 436
DB 357 IAGAASRISGPFSTEDVASNCVPAQSLSEPIIMIVSALMSVGAILLVILLPLFR 416

QY 437 CLRCLRQOHDDFADDISL 455
DB 417 CQR--RPRDPEVNDSESL 433

RESULT 7
Q9JL18 PRELIMINARY; PRT; 514 AA.
AC Q9JL18;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ASPARTYL PROTEASE 1.
GN BACE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Choi D.K., Sugano S., Sakaki Y.;
RT "Molecular characterization of the mouse Aspl gene, a homolog of the
RT human Aspl (down Syndrome Region aspartyl protease).";
RL Submitted (DEC-1999) to the EMBL/Genbank/DBD databases.
DR EMBL; AF216310; AAF36599.1; -.
DR HSSP; P32329; 1YPS.
DR MEROPS; A01.041; -.
DR MGD; MGI:1860440; Bace2.
DR InterPro; IPR001969; Asp_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 3.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 514 AA; 55799 MW; A70725F2C1D5B47 CRC64;

Query Match          47.4%; Score 1146.5; DB 11; Length 514;
Best Local Similarity 50.7%; Pred. No. 2.6e-89;
Matches 216; Conservative 75; Mismatches 130; Indels 5; Gaps 3;

QY 8 EPGR-RGSFVENVDNIRKSGGGYVEMTVGSPPTNLIVDTGSSNFAVGAPHPHLR 66
DB 65 EYVRATANFLAWNDLNGSGRGYLEMLIGTPKQVQLIVDTGSSNFAVGAPHPSTYDT 124

QY 67 YQORQLSSTRDLRKGVYVPTQKWEGLDGLVSTPHGPNVYVRANIIATSDKFKFI 126
DB 125 YFDSESSSTYHSKGFYVYKYTGQSGTGVGSDLVTPKGFNSPLVNIATIFESNFLL 184

QY 127 NGSNMEGILGLAVAEIARPDSDLPEFDSLVRQTHVPMLESLQLCGAGFPLNOSSEVLASV 186
DB 185 PGIKMNGIILGLAVAEIARPDSDLPEFDSLVRQTHVPMLESLQLCGAGFPLNOSSEVLASV 241

QY 187 GGSMTIGGIDHSLYTGSIMWTPPIRREWEYEVIIIVREINGODLMDCKEYNDKSIYVSG 246
DB 242 GGSILVIGIEPSLYKGDIVWTPPIKEWYQIILKLEIGGQNLNLDCKREYNADKAIYVSG 301

QY 247 TTNRLRLPKRVFAAAYKSIKAASTEFKPPGFMGLGBOLYCWOAGTTPMNIPIYISLYLME 306
DB 302 TTNRLRLPKRVFAAAYKSIKAASTEFKPPGFMGLGBOLYCWOAGTTPMNIPIYISLYLME 361

```

```

Oy 307 VTMOSRRTILTPQOYLRPVEDVATSDDDCYKFAISSSGTGWAGVIMSGFIYVDPARK 366
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 362 MNSRSTRITLLPOLYIOPMKGAGFN - ECFRFGISSSTNALVIGATVMEGFVYFPRAOR 420
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 367 RIGFAVSACHVHDEFTTAIVEGPFVLLDMEDCGYNIPQDESLMTIAYMAICALFM 426
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 RVGFVAVSPCAEEGTIVSELSGPFSTIEDIASNCPVPAQALNEPLIMIVSTALMSVCAIILL 480
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 427 PLCLMY 432
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 VLLILL 486
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
O9NZL2
ID O9NZL2 PRELIMINARY; PRT; 468 AA.
AC O9NZL2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE ASPARTYL PROTEASE.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP
RX MEDLINE-20422477; PubMed-10965118;
RX Solans A., Estivill X., de la Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase.";
RL Cytogenet. Cell Genet. 89:177-184(2000).
DR EMBL, AF188276; AAF35835.1; -.
DR HSCP; P00797; ZREN.
DR MEROPS; A01.041; -.
DR InterPro; IPR001969; Asp_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF000026; asp_3.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ
SEQUENCE 468 AA: 50324 MW: 717E0920126A0142 CRC64:

```

	Query Match	40.2%.	Score 972.5;	DB 4;	Length 468,	
	Best Local Similarity	43.1%,	Pred. No. 1.6e-74;			
	Matches 193; Conservative	74;	Mismatches 124;	Indels	57;	Gaps
QY	9	PGRRSFEVEMVDNLGRKGAGGYYVMATGSPQOTNIIIVDGGSSNFPAAGAHPPLHXY	68			
Dd	71	PAGANFLAMVDNLGGDSRGYLEMLIGTPQKIQIILVDGSSNFPAAGTPHSIDTYF	130			
QY	69	QRQLSTYRDLRKGYVYPYTGKKWEGELDTLVSIIPHGNVTVRANIAITESDKFTNG	128			
Dd	131	DTERSTYSRSKGFDPVTKYTQGSMTGFVGEDLVITPKGFNSFLVNIMTFESSENFLPG	190			
QY	129	SNMBEIIQLAELAETLRPDSDLFFPFDLSLYKQTHVNFSLQLCGAPFLNOSEVLASVGG	188			
Dd	191	IKMWIIQLALAVATLAKPSSSLTEFPFDSLVTQTANIPIVNSMOCGGLPVAGS--GTMG	247			
QY	189	SMIIGGDHSLYGTSGLMYTPIRREMYEVIIIVRVIEINODLKMDCKEYNVKYSTVDSCT	248			
Dd	248	SLVIGCIERPSLYKGDIMWTPIKEENYYIIELLKLEIGQSILNDCREYNADKAIVDSST	307			
QY	249	NRLRPKYFEAAVMSIKAASSTERFPDGEFWLGEOLVCMQAQSTPPWNIPVLSYLMEYTV	308			
Dd	308	LRLRPKYFEADVAVARASLL-----	329			
QY	309	NQSRITILLPOOYLARPVEDVATSDDDCYKFAISOSSTGVMGAYIMEGFYVFDAARKRI	368			
Dd	330	-----YIQPMWGACLN-ECYRFISBPNTNALGVATIMEGFYVIFPAOKRV	376			

[illegible]

```

RESULT      9
Q9NZLI
ID          Q9NZLI          PRELIMINARY;      PRT;      396 AA.
AC          Q9NZLI;
DT          01-OCT-2000 (TREMBLrel. 15, Created)
DT          01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT          01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE          ASPARTYL PROTEASE.
GN          BACE2.
OS          Homo sapiens (Human) .
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX          NCBI_TaxID=9606;
LN          [1]
RN          SEQUENCE FROM N.A.
RP          MEDLINE=20422477; PubMed=10965118;
RX          SOLans A., Estivill X., de La Luna S.;
RT          "A new aspartyl protease from 21q22.3, BACE2, is highly similar to
RT          Alzheimer's amyloid precursor protein beta-secretase.";
RL          Cytogenet. Cell Genet. 89:177-184(2000).
DR          EMBL; AF188377; AAF3836.1; -.
DR          HSSP; P00797; 2REN.
DR          MEROPS; A01.041; -.
DR          InterPro; IPR001969; Asp_protease.
DR          InterPro; IPR001461; Pepsin.
DR          Pfam; PFO0026; asp; 2.
DR          PRINTS; PR00792; PEPsin.
DR          PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW          Protease.
SQ          SEQUENCE.      396 AA; 43013 MW; 5023A7AF391CEAC9 CRC64;

```

Query Match	40.0%	Score 967.5	DB 4	Length 396
Best Local Similarity	54.6%	Pred. No. 3.4e-74		
Matches 183	Conservative 54	Mismatches 85	Indels 13	Gaps 4
QY 9	PGRRGSEFVENVNLKSGKSGGQYVYEMTGTGSPQOLINIIIVDTGSSNFAYGAAHPHLHRY	68		
Db	1 PAGAANFLAWYDNLQDGSGRGYLEMLLGTTPQKQIILVDTGSSNFAYAGTHSTIDTYE	130		
QY 69	QROLSTYRDLRKGYVYVYITGCKMEGLDGLVSIPIHPGNVTVRANIAATESDKFFING	128		
Db	131 DFESSSTYRSKGFDPVTKYKTGSSWTFGEEDLVITPIPKGFNTSEIVNITAFISENFPLDG	190		
QY 129	SNMEGIGLAAETARPPDLSLEPFDDSLVKQTHVNLNLSLOLCAGKGFPLNOSVLAISVG	188		
Db	191 IKMNGILGLAATATLAKBPSSLETFDDSLVTOANTIPNVSMQCGAGLPAAS--GTNGG	247		
QY 189	SMIIGGIDHSLYTGSLMWTPIRREMYEVILIVRAYEINQDCKMDCKEYNKDSYDVGST	248		
Db	248 SLVIGGLEPSSLYKGDIMWTPIKEEMYQIILKLEIGQSLINDOREYNADRAKAYDSGTT	307		
QY 249	NLRPKVYEEAAYASIKRAASTEKFPDGFWTGEOGLVCMQAQCTTPWNIIPVLSLYTMEGT	308		
Db	308 LLRLPQKFEVDVAEAVARASLIPEFSQGFMTGSQLACTNSETPWSYPKISITYLREDS	367		
QY 309	NOSFRITILLPOQYLRPEEDVATSQDDCKE--AISO	342		
Db	368 SRSFRITILLPOK-LRVLLQ-----CLKFPGLSQ	393		
RESULT 10				
99PDD2				

```

ID Q9P0D2 PRELIMINARY; PRT: 213 AA.
AC Q9P0D2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HSPC104 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=CORD BLOOD.
RA Zhang Q.H., Ye M., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
RT "Human partial CDS cloned from cd34+ stem cells."
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF161367; AAF28927.1;
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF00026; asp; 1.
FT NON_TER
SQ SEQUENCE 213 AA; 24338 MW; EC9D3FA31CFA835C CRC64;

Query Match 29.5%; Score 712.5; DB 4; Length 213;
Best Local Similarity 83.5%; Pred. No. 7,8e-53;
Matches 137; Conservative 4; Mismatches 12; Indels 11; Gaps 1;

QY 193 GGDHSLYGLSMTYPIREMYEVIIVEINGODLKMCKEYNDKSIIVDSGTTNRL 252
DB 1 GGDHSLYGLSMTYPIREMYEVIIVEINGODLKMCKEYNDKSIIVDSGTTNRL 60

QY 253 PKKVEFAAASKIAASSTKFPDGFGLGBOIYCNQAGTTPMNFPIVLSILMGEVTNOSF 312
DB 61 PKKVEFAAASKIAASSTKFPDGFGLGBOIYCNQAGTTPMNFPIVLSILMGEVTNOSF 120

QY 313 RTIILPQOYLRPEDVATSQDCYKFAISQSSSTGTVMGAVINMEG 356
DB 121 RTIILPQOYLRPEDVATSQDCYKFAISQSSSTGTVMGAVINMEG 153

RESULT 11
Q9R1P7 PRELIMINARY; PRT: 255 AA.
AC Q9R1P7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ASPATYTL PROTEASE (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RC Accarino M., Fumaçalli P., Taramelli R., Ottolenghi S.;
RT "Cloning of a gene from chromosome 21 Down Region encoding a potential
transmembrane protease."
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF051150; AAD45964.1;
DR MEROPS: A01.041;
DR InterPro: IPR001969; Asp_protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF00026; asp; 2.
DR PRINTS: PRO00792; PEPSTN.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
KM PROTEASE.
FT NON_TER
SQ SEQUENCE 255 AA; 28685 MW; 53DE317815996D63 CRC64;

Query Match 24.7%; Score 596.5; DB 11; Length 255;
Best Local Similarity 47.8%; Pred. No. 8e-43;

```

```

Matches 109; Conservative 44; Mismatches 74; Indels 1; Gaps 1;

QY 205 WYTIPIREMYEVIIVEINGODLKMCKEYNDKSIIVDSGTTNRLPKKVEFAAASKI 264
DB 1 WYTIPIREMYEVIIVEINGODLKMCKEYNDKSIIVDSGTTNRLPKKVEFAAASKI 60

QY 265 KAASSTKFPDGFGLGBOIYCNQAGTTPMNFPIVLSILMGEVTNOSFRITLIPQOYLRP 324
DB 61 ARTSILPEFSGFTGAGLACWTSETPMVAFPKISILYLRNENASRSFRITLIPQOYL 120

QY 325 VEDVATSDDCYKFAISQSSSTGTVMGAVINMEGFYVDRARRKRGFVSAACHVDERIA 384
DB 121 MMGAGFNY-ECYRFGISSSTNALVIGATVMEGFYVDRARRRGVAFVSPAELETGTVS 179

QY 385 AVEGFEVTLMDGCGYNIPQDESTLMTIAYVMAICATLFPPLCLMW 432
DB 180 EISGFESTEDIASNCVPAALNEPLMTIVSTALMSVCAILLVILL 227

RESULT 12
Q9VKP6 PRELIMINARY; PRT: 391 AA.
AC Q9VKP6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CG17134 PROTEIN.
CN CG17134.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Suton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelpart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

```


"The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 DR EMBL: AE003630; AAF53016.1; -.
 DR HSSP: P00794; 4CMS.
 DR MEROPS: A01.0FW; -.
 DR F1yase; FBgn0032304; CG17134.
 DR InterPro: IPR001969; Asp_protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp.1
 DR PRINTS: PR00792; PERSIN
 DR PROSITE: PS00141; ASP_PROTEASE; 2.
 DR Aspartyl protease; Hydrolase.
 KW ASSEQUENCE 391 AA; 42016 MW; 99A18E7131025E1B CRC64;
 SQ

Query Match 13.7%; Score 332.5; DB 5; Length 391;
 Best Local Similarity 29.1%; Pred. No. 5e-20;
 Matches 105; Conservative 61; Mismatches 146; Indels 49; Gaps 13;

QY 20 DNLRGKSGGYVEMTVGSPQTLNITVDGSSNFAYGAPHPF-----LHRYQRQLS 73
 DB 66 EHLHSMNNERYGVAIGTPEQRFNILEDTGSANLWBSASCPASNTACQRRNKYDSSAS 125
 QY 74 STYRDLRKGVVVPYTGKWEGLGTLVSIPIGPNVTVRANIAITTESDKFFINGSNMG 133
 DB 126 STYVANGEEFAIEYGTSLSGFLSNDIVTIA-GISIQNGTEALSEFGTTFVD-APRAG 183
 QY 134 ILGLVAELARDDSLPEPFDLSVKQTHVPN-LESLQLCGAGFPLNQSSEVLASVGSMTI 192
 DB 184 ILGLAFSAIA--VDGVTPEDNMISQGLDEPVIISFYLRKG-----TAVRSGEIL 233
 QY 193 GGIDHSILGSLMYPIRREWEYEVIIIVREINQDLMKDCREYNDKSIYDSGTNRL 252
 DB 234 GGIDSLVRSGLTYVPVPAWPKVNTIKTNGLLNCG-----QALDTGSLIAV 287
 QY 233 PKVVEAAVSKSAASSTKEKPPDGFGLGEOLV-CWQAGTTPMNPVLSLYLMEVNTQS 311
 DB 288 PLAAVRKINRGATDND-----GEAFVRCGRVSS-----LPKYNLNIQGV---- 329
 QY 312 FRITLPOOYLRPVEDVATSDDDCK-PAISQSSGTVMGAVIMGEFVVDRAKRRIGF 370
 DB 330 --FTLAPRDYI--KVYTONGQYCKSAFTYMEGLSEFWILGDFYIGKFTYVDKNERIGF 385
 QY 371 A 371
 DB 386 A 386

RESULT 13
 Q9GYX7 PRELIMINARY; PRT; 354 AA.
 AC Q9GYX7;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 OS HEME-BINDING ASPARTIC PROTEINASE (FRAGMENT).
 CC Boophilus microplus (Cattle tick).
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
 CC Parasitiformes; Ixodidae; Ixodidae; Boophilus.
 CC NCBI_Taxid=6941;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PORTO ALEGRE; TISSUE=OVARY;
 RA Sorgiano M.H.F., Loguillo C., Zingali R.B., Paiva-Silva G.O.,
 RA Juliano L., Oliveira P.L.;
 RT "A heme-binding aspartic proteinase from the eggs of the hard tick
 RT Boophilus microplus".
 RL J. Biol. Chem. 0:0-0(2000).
 DR EMBL: AF286865; AAG00993.1; -.
 DR HSSP: P00797; 2REN.
 DR MEROPS: A01.054; -.
 SQ

DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp.1.
 DR PRINTS: PR00792; PERSIN.
 FT NON_TER 1
 SQ SEQUENCE 354 AA; 38512 MW; B50D7C638CF27091 CRC64;
 SQ

Query Match 13.7%; Score 332; DB 5; Length 354;
 Best Local Similarity 26.2%; Pred. No. 4.8e-20;
 Matches 101; Conservative 74; Mismatches 135; Indels 76; Gaps 15;

QY 19 VDNLRGK-----SGQGYVEMTVGSPQTLNITVDGSSNFAYGAPHPFL- 64
 DB 13 VTEIRGRLSDPIPIILTYNNMOPFGITIGTPQSKFLMDGSSNFVW-----PSIN 67
 QY 65 -----HRYQRQLSSTYRDLRKGVVVPYTGKWEGLGTLVSIPIGPNVTVRANI 115
 DB 68 CDQSMACRHDHAKYDSKSTFTKSGRYIRIRYSGVYRGISIDNVCV--GPATVQYKF 125
 QY 116 AAITESDKFFINGSWESITLALAEIARPDLSLEPFDLSVKQTHVPN-LESLQLCGAG 174
 DB 126 AEMHSDGKLFERNAKYDIFGLAPPSISQ--NQLPLFDAMVKGQVROAVFSIYL--SK 181
 QY 175 FPLNQSEVLASVGSMTIIGIDHSILGSLMYPIRREWEYEVIIIVREINQDLMK-DC 233
 DB 182 QPSDGN-----GGEIYEGGINAQRITGAIHYVVSQAAMQVMDINVOGTTLVGGC 235
 QY 234 KEYNDKSIYDSGTNTRLRPKVFEAAVKSIAASSTKEKPPDGFGLGEOLV-CWQAGTTPW 293
 DB 236 -----PTVDSGTSLSGP-----SADVETLNRVIGATKTAAG-----FEVNCATI 277
 QY 294 NIPVVISLYLME---VTNOSFRITILPQOYLRPVEDVATSDDDCYFAISQSSGT--- 347
 DB 278 SSLPPIITFNLGKSPFIQGEAYTIRI-----PLTGGEGCFTRISEDSASGTNLW 327
 QY 348 VMGAVIMEGEFVVDRAKRRIGFVAVS 373
 DB 328 ILGAVFTQTYTYVDFRAQNRRGFATA 353

RESULT 14
 Q9BGUS PRELIMINARY; PRT; 386 AA.
 AC Q9BGUS;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 OS CATHEPSIN D (FRAGMENT).
 CC CAT-D.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Higuchi M., Miyashita N., Nagamine Y., Awata T.;
 RT "Complementary DNA sequence of bovine cathepsin D.";
 RT Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 CC EMBL: AB055312; BAB21620.1; -.
 DR HSSP: P07339; 1LYB.
 DR MEROPS: A01.009; -.
 DR InterPro: IPR001969; Asp_protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp.1.
 DR PRINTS: PR00792; PERSIN.
 DR PROSITE: PS00141; ASP_PROTEASE; 2.
 KW Aspartyl protease; Hydrolase.
 FT NON_TER 1
 SQ SEQUENCE 386 AA; 41965 MW; 33BF00B80585490D CRC64;
 SQ

Wed Oct 30 14:11:50 2002

us-09-724-571-43.rsp

Page 8

Query Match	12.9%	Score 312.5;	DB 6;	Length 386;
Best Local Similarity	28.0%;	Pred. No. 2.5e-18;		
Matches 113;	Conservative	69;	Mismatches 131;	Indels 91;
				Gaps 20;

[illegible]

```

RESULT      15
ID          09DEC2
AC          09DEC2      PRELIMINARY;      PRT;      384 AA.
DT          01-MAR-2001 (TREMBLrel. 16, Created)
DT          01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT          01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE          PEPsinogen A.
OS          Xenopus laevis (African clawed frog).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoae; Pipidae;
OC          Xenopodinae; Xenopus.
OX          NCBI_TaxID=8335;
RN          [1]
RP          SEQUENCE FROM N.A.
RA          Ikusawa M., Inokuchi T., Kobayashi K., Yasumasa S.;
RT          "Molecular Cloning of Pepsinogens in Adult Xenopus laevis and Bullfrog
RT          Rana catesbeiana."
RT          Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
CC          -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC          EUKARYOTIC ASPARTYL PROTEASES FAMILY.
DR          EMBL: AB045380; BAB20798.1; -.
DR          HSSP; P00790; IPSN.
DR          MEROPS; A01.001; -.
DR          InterPro: IPR001669; Asp_protease.
DR          InterPro: IPR001461; Pepsin.
DR          Pfam: PF00026; asp. 1.
DR          PRINTS: PR00792; PEPsin.
DR          PROSITE: PS00141; ASP_PROTEASE; 1.
KW          Aspartyl protease; hydrolase.
SQ          SEQUENCE 384 AA; 4188 MW; 729DD2E7933D9072 CRC64;

```

	Query Match	12.7%	Score 308;	DB 13;	Length 384;
	Best Local Similarity	27.5%;	Pred. No. 66-18;		
	Matches	99;	Mismatches 133;	Indels 68;	Gaps 16.
QY	30 YYYEATVGSPPQTLNLIVDTGSSNFVAGAPHFLL-----HRYQRQLSTSYRLRLKRG	82			
	: : : : : : : :				

```

Db      72  YGTTISICTPPOEFTVTFIDTGSANLW---PSYVCSSQACSNHNRFPQOOSTFOATNTP 128
QY      83  VYVPYTGKRWEGELGTDLVSIPIGPNTVTRANIAITBESDK-PIFINGSNNEGILGLAYAE 141
Db      129  VSIQGTGSMGSLGAYDPLTQV---GNIOISNOMGSLSEBGSFLYSPEDGILGLAFPS 185
QY      142  IARPDGSLPEPFDSLVKQTHVP-NLEFSLOLGGAGFPLINOSEVLASVGSAMIIGIDHSLY 200
Db      166  IIA--SSQATPVFDMNMGSLIPQMLFSEVYLSSDG-----QTSYVLEFGVDNSY 233
QY      201  TGSMTVPIRREMYEVLIVAEINQDL--KMOCKEYNDKSIYDGGTNTLRPKKVE 258
Db      234  SGSLMWVPLAEYVQITLDSVTSINGOVLACQSC-----QALVDTGTSIMTGPSTPI- 266
QY      259  AAVRSIKKAASSTREKPPDFEWMGEOLV-CWAQGTTPWNIFFVPLISLYLGEVYNOSFRITIL 317
Db      287  ANIQVYIGASQDSN-----GQVYINONNINSNMPTIIF-----TIN 321
QY      318  PQQY-LRPVEDATNSQDCKY-FALSQSTGT-----WGVATIMEGFVYEDRARKRIGFA 371
Db      322  GVQYPLSPSAIVKRONQCGSSSFQAMNLPJNSGLMTLIGDVFIRQYTFVDFBRANNYVAIA 384

```

Search completed: August 7, 2002, 09:17:12
Job time: 216 sec

381

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2002, 09:13:36 ; Search time 17.62 Seconds
(without alignments)
1002.050 Million cell updates/sec

Title: US-09-724-571-43

Perfect score: 2419
Sequence: 1 ETDEPEPGRGSGFVEMVD.....CLRLRQHHDFADISILK 456

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2419	100.0	501	1	BACE_HUMAN
2	2392	98.9	501	1	BACE_RAT
3	2390	98.8	501	1	BACE_MOUSE
4	1171.5	48.4	518	1	BACE2_HUMAN
5	327	13.5	324	1	PEP1_GADMO
6	314.5	13.0	390	1	CATD_BOVIN
7	306	12.6	387	1	PEP1_RABIT
8	302.5	12.5	388	1	PEP4_MACRU
9	302	12.5	367	1	PEP4_CHICK
10	299.5	12.4	366	1	CATE_HUMAN
11	298.5	12.3	383	1	PEPE_CHICK
12	296	12.2	387	1	PEP2_RABIT
13	296	12.2	391	1	CATE_CAVPO
14	295.5	12.2	412	1	CATD_HUMAN
15	295	12.2	387	1	PEP4_RABIT
16	289.5	12.0	388	1	PEP2_MACRU
17	287	11.9	407	1	CATD_RAT
18	286	11.8	387	1	PEP3_RABIT
19	285.5	11.8	398	1	CATE_RAT
20	285	11.7	388	1	CATD_CHICK
21	283.5	11.7	388	1	PEP4_MACMU
22	283.5	11.6	381	1	CHYM_SHEEP
23	281.5	11.6	388	1	PEP1_MACRU
24	281.5	11.6	410	1	CATD_MOUSE
25	279.5	11.5	386	1	PEP4_PIG
26	278.5	11.5	387	1	PEP4_CALJA
27	278	11.5	396	1	CATE_RABIT
28	277.5	11.5	397	1	CATE_MOUSE
29	277.5	11.4	419	1	CARV_CAVNL
30	276.5	11.4	381	1	CHYM_BOVIN
31	273.5	11.3	381	1	PEPF_RABIT
32	273	11.3	377	1	PEPC_MACRU
33	270.5	11.2			

ALIGNMENTS

RESULT ID	1	BACE_HUMAN	STANDARD	PRT	501 AA
AC	P56817	Q9UT5			
DT	30-MAY-2000	(Rel. 39, Created)			
DT	30-MAY-2000	(Rel. 39, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)				
DE	Beta-site amyloid precursor protein cleaving enzyme (Aspartyl				
DE	protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)				
DE	(Memapsin-2).				
GN	BACE OR BACE1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_Taxid=9606;				
RP	[1]	SEQUENCE FROM N.A. (ISOFORM A).			
RP	TISSUE=Brain;				
RC	MEDLINE=20002972; PubMed=10531052;				
RA	Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,				
RA	Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,				
RA	Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,				
RA	Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,				
RT	Tianrui J., Rogers G., Cilton M.;				
RT	"Beta-secretase cleavage of Alzheimer's amyloid precursor protein by				
RL	the transmembrane aspartic protease BACE.";				
RL	Science 286:735-741(1999).				
RP	[2]	SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 46-68, AND			
RP	CHARACTERIZATION.				
RC	TISSUE=Brain;				
RC	MEDLINE=20057171; PubMed=10591214;				
RA	Silva S., Anderson J.P., Barbour R., Basl G.S., Caccavello R.,				
RA	Davis D., Dean M., Dovey H.F., Frigon N., Hong J., Jacobson-Croak K.,				
RA	Jewett N., Keim P., Knops J., Lieberburg I., Power M., Tan H.,				
RA	Takano G., Tung J., Schenk D., Seubert P., Simonsen S.M., Wang S.,				
RA	Walker D., Zhao J., McConlogue L., Varghese J.;				
RT	"Purification and cloning of amyloid precursor protein beta-secretase				
RT	from human brain.";				
RL	Nature 402:537-540(1999).				
RP	[3]	SEQUENCE FROM N.A. (ISOFORM A).			
RC	MEDLINE=20057170; PubMed=10591213;				
RA	Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,				
RA	Brashers J.R., Strattan N.C., Mathews W.R., Buhl A.E., Carter D.B.,				
RA	Tomaselli A.G., Parodi L.A., Helinikson R.L., Gurney M.E.,				
RT	"Membrane-anchored aspartyl protease with Alzheimer's disease beta-				
RT	secretase activity.";				
RL	Nature 402:533-537(1999).				
RP	[4]	SEQUENCE FROM N.A. (ISOFORM A).			
RC	MEDLINE=20030166; PubMed=10561122;				
RA	Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,				
RA	Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,				
RA	Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;				

RT "Identification of a novel aspartic proteinase (Asp 2) as beta-secretase.";
 RL Mol. Cell. Neurosci. 14:419-427(1999).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM B).
 RC TISSUE=Brain, and Pancreas;
 RA Michel B., De Pietri Tonelli D., Zacchetti D., Keller P.;
 RT "New beta-site APP cleaving enzyme isoform (BACE-1B) obtained from human brain and pancreas.";
 RU Submitted (Jan-2001) to the EMBL/Genbank/DBJ databases.
 RP [6]
 RP SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.
 RA MEDLINE=20144060; PubMed=10677483;
 RL Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein.";
 RU Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
 CC -1- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP, LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; A/BACE-1A (shown here) and B/BACE-1B; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF190725; AAF04142.1; -;
 DR EMBL; AF201468; AAF18982.1; -;
 DR EMBL; AF200343; AAF17079.1; -;
 DR EMBL; AF204943; AAF26367.1; -;
 DR EMBL; AF338816; AAK38374.1; -;
 DR EMBL; AF200193; AAF13715.1; -;
 DR HSSP; P32329; IYPS.
 DR MIM; 604252; -;
 DR InterPro; IPR001969; Asp.-protease.
 DR InterPro; IPR001461; Pepsin.
 DR Pfam; PF00026; asp; 3.
 DR PRINTS; PR00792; PEPsin.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 KW Hydrolyase; Aspartyl protease; Glycoprotein; zymogen; Transmembrane;
 KW Signal; Alternative splicing.
 FT SIGNAL 1 21
 FT PROPEP 22 45
 FT CHAIN 46 501
 FT DOMAIN 22 457
 FT TRANSMEM 458 478
 FT DOMAIN 479 501
 FT ACT_SITE 93 93
 FT ACT_SITE 289 289
 FT CARBOHYD 153 153
 FT CARBOHYD 172 172
 FT CARBOHYD 223 223
 FT CARBOHYD 354 354
 FT VARSPLIC 190 214
 SQ SEQUENCE 501 AA; 55763 MW; 377CE4C824ACEF05 CRC64;

Query Match 100.0%; Score 2419; DB 1; Length 501;
 Best Local Similarity 100.0%; Pred. No. 5, 3e-193;
 Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEPEPGRGRGSFVENVNDLRKSGSGYVYEMTVGSPQTLNIIIVDTGSSNFAVGAAP 60
 Db 46 ETDEPEPGRGRGSFVENVNDLRKSGSGYVYEMTVGSPQTLNIIIVDTGSSNFAVGAAP 105
 QY 61 HPFLHRYQROLSSTYRDLRKGVVYPYQKMEGELGTLVSIHPGNVTVRANIAATE 120
 Db 106 HPFLHRYQROLSSTYRDLRKGVVYPYQKMEGELGTLVSIHPGNVTVRANIAATE 165
 QY 121 SDKEFINSNMEGTLGLAETIARPDLSLEFPFSLVKQTHVPLVLSQLCGAGFPINOS 180
 Db 166 SDKEFINSNMEGTLGLAETIARPDLSLEFPFSLVKQTHVPLVLSQLCGAGFPINOS 225
 QY 161 EYLVSGVSMIIGIDHSLYTGSLMYTPIRREWYEVITVVEINIGDGLKMDCKEYNDK 240
 Db 226 EYLVSGVSMIIGIDHSLYTGSLMYTPIRREWYEVITVVEINIGDGLKMDCKEYNDK 285
 QY 241 STVDGTTNLRPKKVFPAVAKSIKAASSTKFPDGFMLGQLYCWMAGTTPNNIFPVIS 300
 Db 286 STVDGTTNLRPKKVFPAVAKSIKAASSTKFPDGFMLGQLYCWMAGTTPNNIFPVIS 345
 QY 301 LYLMEVYNOSFRITLLPQOYLRPVEDVATSDODCYKFAISOSTGTVMGAVIMEGFYV 360
 Db 346 LYLMEVYNOSFRITLLPQOYLRPVEDVATSDODCYKFAISOSTGTVMGAVIMEGFYV 405
 QY 361 FDRARRIGFPAVSACHVDEPFAVEGPFVTLDMEDGYNIPQDESTLMTIAYVMAI 420
 Db 406 FDRARRIGFPAVSACHVDEPFAVEGPFVTLDMEDGYNIPQDESTLMTIAYVMAI 465
 QY 421 CALFMLPLCLAWCQWRCLRCRQOHDFADISLILK 456
 Db 466 CALFMLPLCLAWCQWRCLRCRQOHDFADISLILK 501
 RESULT 2
 ID BACE_RAT STANDARD: PRT; 501 AA.
 AC P56819;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
 DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
 DE protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)
 DE (Memapsin-2).
 GN BACE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20002972; PubMed=10531052;
 RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
 RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
 RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
 RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
 RA Treanor J., Rogers G., Citron M.;
 RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.";
 RU Science 286:735-741(1999).
 CC -1- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP, LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF190727; AAF04144.1; -.
 DR HSSP: P32329; 1YPS.
 DR InterPro: IPR001969; Asp_protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp; 3.
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE; 1.
 KM Hydrolyse; Aspartyl protease; Zymogen; Transmembrane;
 KW Signal.
 FT SIGNAL. 1 21 POTENTIAL.
 FT PROPEP 22 45 POTENTIAL.
 FT CHAIN 46 501 BETA-SECRETASE.
 FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 458 478 POTENTIAL.
 FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
 FT ACT_SITE 93 93 BY SIMILARITY.
 FT ACT_SITE 289 289 BY SIMILARITY.
 FT CARBOHD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 501 AA; 55806 MW; 24B445BC8BE87DE3 CRC64;

Query Match 98.9%; Score 2392; DB 1; Length 501;
 Best Local Similarity 98.2%; Pred. No. 9.2e-191;
 Matches 448; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 ETDEPEEPGRGRSGFVFNVDNLGRKSGGGYVEMTGSPPQTLNLTIVDTGSSNFAVGAAP 60
 DB 46 ETDEPEEPGRGRSGFVFNVDNLGRKSGGGYVEMTGSPPQTLNLTIVDTGSSNFAVGAAP 105
 QY 61 HPEFLRRYQROLSSTYRDLRKGVYPRYQGWGELGTDLVSIHPGPNVTVRANATAITE 120
 DB 106 HPEFLRRYQROLSSTYRDLRKGVYPRYQGWGELGTDLVSIHPGPNVTVRANATAITE 165
 QY 121 SDKFFINSNMGGILGLVAETARPDDSLPEPDSLVKOTHPNLFSLQLOGAGFPNWS 180
 DB 166 SDKFFINSNMGGILGLVAETARPDDSLPEPDSLVKOTHPNLFSLQLOGAGFPNWS 225
 QY 181 EVLASVSGSMIIGGIDHSLYTGSMTYTPIRREMYEVIIVREINIGODLKMCKEYNYDK 240
 DB 226 EVLASVSGSMIIGGIDHSLYTGSMTYTPIRREMYEVIIVREINIGODLKMCKEYNYDK 285
 QY 241 SIYDSGTINLRPKKVFENAAVSKTAASSTKFPDGFVLGEOLVCMQAGTTPWNIFFVIS 300
 DB 286 SIYDSGTINLRPKKVFENAAVSKTAASSTKFPDGFVLGEOLVCMQAGTTPWNIFFVIS 345
 QY 301 LYLMGEVNTNOSFRITILPOQYLRPEVATSDODCYKFAISOSSGTMGAVIMGEFVV 360
 DB 346 LYLMGEVNTNOSFRITILPOQYLRPEVATSDODCYKFAISOSSGTMGAVIMGEFVV 405
 QY 361 FDRARRKIRGFANVACHVDEFTAAVEGPFYTLDMEDGYNIPQDESTLMTIAVMAAI 420
 DB 406 FDRARRKIRGFANVACHVDEFTAAVEGPFYTLDMEDGYNIPQDESTLMTIAVMAAI 465
 QY 421 CALFPLPCLLWVCQRCRLCRQOHDDPADDISLTK 456
 DB 466 CALFPLPCLLWVCQRCRLCRQOHDDPADDISLTK 501

RESULT 3
 BACE_MOUSE STANDARD; PRT; 501 AA.
 AC P56818;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
 DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
 DE protease 2) (Asp 2) (Asp2) (Membrane-associated aspartic protease 2)
 DE (Memapsin 2).
 CN BACE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20002972; PubMed=10531052;
 RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
 RA Denis P., Teplow D.B., Ross S., Amarante P., Loeflof R., Luo Y.,
 RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
 RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
 RA Treanor J., Rogers G., Citron M.;
 RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
 RT the transmembrane aspartic protease BACE.";
 RL Science 286:735-741(1999).
 RN [2]
 RP REVISIONS TO 6 AND 81-87.
 RA Bennett B.D., Vassar R., Citron M.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057170; PubMed=10591213;
 RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
 RA Brashler J.R., Strachan N.C., Mathews W.R., Buhl A.E., Carter D.B.,
 RA Tomasselli A.G., Parodi L.A., Heinrichson R.L., Guney M.E.;
 RT "Membrane-anchored aspartyl protease with Alzheimer's disease
 RT beta-secretase activity.";
 RL Nature 402:533-537(1999).
 CC -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
 CC AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
 CC THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,
 CC LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
 CC SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
 CC FRAGMENT WHICH IS LATER RELEASE BY GAMMA-SECRETASE (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: BRAIN.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF190726; AAF04143.2; -.
 DR EMBL: AF200346; AAF17082.1; -.
 DR HSSP: P56272; 1AM5.
 DR MGD: MGI:1346542; Bace.
 DR InterPro: IPR001969; Asp_protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp; 1.
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE; 1.
 KW Hydrolyse; Aspartyl protease; Zymogen; Transmembrane;
 KW Signal.
 FT SIGNAL. 1 21 POTENTIAL.
 FT PROPEP 22 45 POTENTIAL.
 FT CHAIN 46 501 BETA-SECRETASE.
 FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 458 478 POTENTIAL.
 FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
 FT ACT_SITE 93 93 BY SIMILARITY.
 FT ACT_SITE 289 289 BY SIMILARITY.

```

FT CARBOHD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 501 AA; 55747 MW; C085A013145E474E CRC64;

Query Match 98.8%; Score 2390; DB 1; Length 501;
Best Local Similarity 98.2%; Pred. No. 1.3e-190;
Matches 448; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 EDDPEEPGRGSRGSEVEMVDNLKRGSGGYYEMTVGSPPTNLITLVDTGSSNFAVGAAP 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 46 ETDSESEBGRGSRGSEVEMVDNLKRGSGGYYEMTVGSPPTNLITLVDTGSSNFAVGAAP 105

OY 61 HPELHRYOROLSTYRDLRKGVVYPTOGKWEGLGTDVLSIPGPNVTYRANIAAITE 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 106 HPELHRYOROLSTYRDLRKGVVYPTOGKWEGLGTDVLSIPGPNVTYRANIAAITE 165

OY 121 SDFKFTNGSNWGIIIGLAFAELIARPDSDLEPFDSLVKOTHPNLFSIOLGAGFPLNOS 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 166 SDFKFTNGSNWGIIIGLAFAELIARPDSDLEPFDSLVKOTHPNLFSIOLGAGFPLNOS 225

OY 181 EYLAIVSGMIIIGIDHSIYTGSLWYTPIRREMYEVIIVREINGODLKMCKEYNYDK 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 226 EYLAIVSGMIIIGIDHSIYTGSLWYTPIRREMYEVIIVREINGODLKMCKEYNYDK 285

OY 241 SYVDSGTTLRLPKVFEAAVKSIAASTEKPPDGFMLGEOVLVQOAGTTWNTFFPVIS 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 286 SYVDSGTTLRLPKVFEAAVKSIAASTEKPPDGFMLGEOVLVQOAGTTWNTFFPVIS 345

OY 301 LYMGVTVQNSRITITLPOYLRPVEDVATSDDDCYKFAISOSSTGTVGAVIMEGFYV 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 346 LYMGVTVQNSRITITLPOYLRPVEDVATSDDDCYKFAISOSSTGTVGAVIMEGFYV 405

OY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDESLTMTIAYMAAI 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 406 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDESLTMTIAYMAAI 465

OY 421 CALFMLPLCLMWCOMRCLRCLROQHDHDFADDSLILK 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 466 CALFMLPLCLMWCOMRCLRCLROQHDHDFADDSLILK 501

RESULT 4
BAE2_HUMAN STANDARD; PRT; 518 AA.
ID BAE2_HUMAN STANDARD; PRT; 518 AA.
AC Q9Y5Z0; Q9YJTG; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta secretase 2 precursor (EC 3.4.23.-) (Beta-site APP-cleaving
DE enzyme 2) (Aspartyl protease 1) (Asp 1) (Asp1) (Membrane-associated
DE aspartic protease 1) (Memapsin-1).
GN BACE2 OR ASP21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;

RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=2005170; PubMed=10591213;
RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA Tomasselli A.G., Parodi L.A., Heinrichson R.L., Gurney M.E.;
RA "Membrane-anchored aspartyl protease with Alzheimer's disease
RT beta-secretase activity.";
RT Nature 402:533-537(1999).
RN [2]
RN SEQUENCE FROM N.A.
RX TISSUE=bone marrow;
RA Xin H., Stephens J.C., Duan X., Harrowe G., Kim E., Grieshammer U.,
RA Giese K.;

```

```

RT "Identification of a novel aspartic-like protease differentially
RT expressed in human breast cancer cell lines.";
RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Accarino M.P., Fumagalli P., Ottolenghi S., Taramelli R.;
RT "Cloning of a gene from chromosome 21 Down region encoding a potential
RT transmembrane aspartyl protease.";
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Solans A., Estivill X., de la Luna S.;
RT "Cloning of a novel mammalian aspartyl protease.";
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RX MEDLINE=20030166; PubMed=10561122;
RA Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
RA Smith T.S., Simons D.L., Walsh F.S., Dingwall C., Christie G.;
RT "Identification of a novel aspartic proteinase (Asp 2) as
RT beta-secretase.";
RT Mol. Cell. Neurosci. 14:419-427(1999).
RN [6]
RN SEQUENCE FROM N.A.
RX MEDLINE=20144060; PubMed=10677483;
RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of
RT proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
RN [7]
RN SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Onki K., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RT Nature 405:311-319(2000).
RN [8]
RN SUBCELLULAR LOCATION: Type I membrane protein.
RN [9]
RN SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
RN EUKARYOTIC ASPARTYL PROTEASES FAMILY.
RN [10]
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN [11]
RN EMBL: AF200342; AAF17078.1; -
RN EMBL: AF117892; AAD45240.1; -
RN EMBL: AF050171; AAD45963.1; -
RN EMBL: AF178532; AAF29494.1; -
RN EMBL: AF204944; AAF26368.1; -
RN EMBL: AF200192; AAF13714.1; -
RN EMBL: AL163284; CAB90458.1; -
RN EMBL: AL163285; CAB90554.1; -
RN HSSP: P00797; 2REN.
RN MIM: 605668; -
RN InterPro: IPR001969; Asp_protease.
RN InterPro: IPR001461; Pepsin.
RN Pfam: PF00026; asp; 3.
RN PRINTS: PR00792; PEPsin.

```

[illegible]

```

RL  PERS Lett. 319:54-58(1993).
RN  [2]
RP  SEQUENCE OF 45-390, AND X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RC  TISSUE=Liver;
RX  MEDLINE=93223670; PubMed=8467789;
RA  Metcalf P, Fusek M;
RT  "Two crystal structures for cathepsin D: the lysosomal targeting
    signal and active site.";
RL  EMO J. 12:1293-1302(1993).
CC  -1- FUNCTION: CATHEPSIN D IS AN ACID PROTEASE ACTIVE IN INTRACELLULAR
    PROTEIN BREAKDOWN.
CC  -1- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
    that of pepsin A. Does not cleave the 4-Gln-1-His-5 bond in B
    chain of insulin.
CC  -1- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC  -1- SUBCELLULAR LOCATION: Lysosomal.
CC  -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
    EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC  PIR: S32383; S32383.
DR  PIR: S37419; S37419.
DR  HSSP: P07339; 1L1B.
DR  MEROPS: A01.009; -.
DR  InterPro: IPR001969; Asp_protease.
DR  Pfam: PF00026; asp. 1.
DR  InterPro: IPR001461; Pepsin.
DR  PRINTS: PR00792; Pepsin.
DR  PROSITE: PS00141; ASP_PROTEASE; 2
    Hydrolyase; Aspartyl protease; Glycoprotein; Lysosome; Zymogen.
FT  PROPEP 1 44
FT  CHAIN 45 390
FT  ACT_SITE 77 77
FT  ACT_SITE 77 77
FT  ACT_SITE 77 77
FT  DISULFID 90 140
FT  DISULFID 90 140
FT  DISULFID 264 268
FT  DISULFID 307 344
FT  CARBOHYD 114 114
FT  CARBOHYD 241 241
FT  SEQUENCE 390 AA; 42488 MW; 5B38AA1C3C48D35 CRC64;
SQ

```

```

RESULT 7
PEP1_RABIT ID PEP1_RABIT STANDARD; PRT; 387 AA.
AC P28712; 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pepsin II-1 precursor (EC 3.4.23.1) (Pepsin A).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91009127; PubMed=2129536;
RA Kageyama T., Tanabe K., Koike O.;
RT "Structure and development of rabbit pepsinogens. Stage-specific
    zymogens, nucleotide sequences of cDNAs, molecular evolution, and
    gene expression during development.";
RL J. Biol. Chem. 265:17031-17038(1990).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS
    INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
    ALSO CLEAVED TO SOME EXTENT.
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE PHE-, LEU-.
CC -1- DEVELOPMENTAL STAGE: PEPINOGENS IN GROUP I, II, AND III WHERE
    THE PREDOMINANT ZYMOGENS AT LATE POSTNATAL STAGE.
CC -1- MISCELLANEOUS: THE EXPRESSION OF PEPINOGEN GENES IS REGULATED BY
    HORMONES AND RELATED SUBSTANCES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
    EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC  PIR: B38302; B38302.
DR  HSSP: P00791; 1PSA.
DR  MEROPS: A01.001; -.
DR  InterPro: IPR001969; Asp_protease.
DR  InterPro: IPR001461; Pepsin.
DR  Pfam: PF00026; asp. 1.
DR  PRINTS: PR00792; Pepsin.
DR  PROSITE: PS00141; ASP_PROTEASE; 2
    Hydrolyase; Aspartyl protease; Digestion; Stomach; Gastric juice;
    zymogen; signal; Phosphorylation; Multigene family.
FT  SIGNAL 1 15
FT  PROPEP 16 59
FT  CHAIN 60 387
FT  MOD_RES 129 129
FT  ACT_SITE 93 93
FT  ACT_SITE 276 276
FT  DISULFID 106 111
FT  DISULFID 267 271
FT  DISULFID 310 343
FT  SEQUENCE 387 AA; 42070 MW; A6EC48F715541A48 CRC64;
SQ

```

```

Query Match 13.0%; Score 314.5; DB 1; Length 390;
Best Local Similarity 28.0%; Pred. No. 1.3e-18;
Matches 113; Conservative 72; Mismatches 128; Indels 91; Gaps 21;

```

```

Query Match 12.6%; Score 306; DB 1; Length 387;
Best Local Similarity 27.1%; Pred. No. 6.4e-18;
Matches 98; Conservative 67; Mismatches 131; Indels 66; Gaps 15;

```


QY 260 AVSINKASSTKPPGFWLGEOLV-CWQAGTTPWNIPEVISTYLMEVTVNOSFRITILP 318
DB 292 IQSYIGASKNL-----LGENIISCSAIDSLPDIYF-----TINN 325
QY 319 QQTLRPVED-VATSDODC---YKFAISQSSTGT--WGVAVIMEGFVYEDRARRKRGIFAV 372
DB 326 VQVPLPASAIILKEDDDCLSGFDGMNLDIVSYGELWILGDVFIQHYFTVFDPRANNQVGLAA 385
QY 373 SA 374
DB 386 AA 387
RESULT 8
PEP4_MACFU STANDARD; PRT: 388 AA.
ID PEP4_MACFU
AC P27678;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Pepsin A-4 precursor (EC 3.4.23.1) (Pepsin I/II).
GN PGA.
OS Macaca fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
NCBI_TaxID=9543;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-70.
RC TISSUE=gastric mucosa;
RX MEDLINE=92037645; PubMed=1935977;
RA Kagiyama T., Tanabe K., Kolwai O.;
RT "Development-dependent expression of isozymes of monkey
RL Eur. J. Biochem. 202:205-215(1991).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY, ALTHOUGH BONDS
CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC ALSO CLEAVED TO SOME EXTENT.
CC -1- CATALYTIC ACTIVITY: PREPERENTIAL CLEAVAGE PHE-, LEU-.
CC -1- MISCELLANEOUS: THE EXPRESSION OF PEPINOGEN GENES IS REGULATED BY
CC HORMONES AND RELATED SUBSTANCES.
CC -1- MISCELLANEOUS: EACH PEPSINOGEN IS CONVERTED TO CORRESPONDING
CC PEPSIN AT PH 2.0 IN PART AS A RESULT OF THE RELEASE OF A 47 AA
CC ACTIVATION SEGMENT AND IN PART AS A RESULT OF STEWISE PROTEOLYTIC
CC CLEAVAGE VIA AN INTERMEDIATE FORM(S).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59753; CAA42425.1; -
DR PIR: S16065; S16065.
DR PIR: S19682; S19682.
DR HSSP: P00790; IPSN.
DR MEROPS: A01.001; -;
DR InterPro: IPR001969; Asp_protease.
DR InterPro: IPR001791; Laminin_G.
DR Pfam: PF00026; asp.1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Digestion; Stomach; Gastric juice;
FT SIGNAL 1 15
FT PROPEP 16 38 BY SIMILARITY
FT PROPEP 39 62 ACTIVATION PEPTIDE.
FT CHAIN 63 388 PEPSIN A-4.

FT ACT_SITE 94 94 BY SIMILARITY.
FT ACT_SITE 277 277 BY SIMILARITY.
FT DISULFID 107 112 BY SIMILARITY.
FT DISULFID 268 272 BY SIMILARITY.
FT DISULFID 311 344 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC...);
SQ SEQUENCE 388 AA; 41955 MW; A2923AB1F7FCEDB9 CRC64;
Query Match 12.5%; Score 302.5; DB 1; Length 388;
Best Local Similarity 27.6%; Pred. No. 1.3e-17;
Matches 107; Conservative 64; Mismatches 133; Indels 83; Gaps 17;
QY 3 DEEPPEGRGRGSPFVEMVDNLKRGSSGGYVEMTVGSPPTGLNLDVTGSSNFAVGAAPHP 62
DB 64 DEQPLE-----NLIDV-----EYFGTIGIGTPAONFTVVEPDTGSSNLMV--PSV 105
QY 63 FL-----HRYVQRLSSTYRDLRKGVVVPYTOGKEGELGTDLSIPHGNVTVRANI 115
DB 106 YCVSLACMDHNLFPQDSSYVATSKTVSITYGTGSMGILGTDYKV--GGISPTNOI 162
QY 116 AATTESDK-FYINGSNWEGILGLAVELARPDSLEPFDSLVKQTHV-PNLFSLQLCGA 173
DB 163 FGLSETEPEGFIFYFAFFDGLGLAVPSIS--SSGATPVEDINQRLVQDLFSVYLSAD 220
QY 174 GFLPNOSEVLASVGSMLIGIDHSLYGSLWYTPIRREMYEVILVREINQDL--KM 231
DB 221 ----DOS-----GSVYLFGLGIDSSYTGSLWVPSVSGVWQISVDSITNMGKTACAK 270
QY 232 DKEKNYDKSIVDSGTTNLRPKVFEAAVKSINKASSTKPPGFWLGEOLV-CWQAGT 290
DB 271 GC-----QAIYDGTGSLTGTPTSPANIQSDIGASNSD-----GEMVSSSAISS 316
QY 291 TFWNIPEVISTYLMEVTVNOSFRITILPQY-LRVEDVATSDQYK-----FAISQSS 344
DB 317 LDIYF-----TINGVQYPLPPSAVILQSGSGCTSGFQGMVPTESG 358
QY 345 TGTGAVIMEGFVYEDRARRKIGFA 371
DB 359 ELWILGDVIRQYFTVFDPRANNQVGLA 385
RESULT 9
PEP4_CHICK STANDARD; PRT: 367 AA.
ID PEP4_CHICK
AC P00793;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pepsin A precursor (EC 3.4.23.1).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=84004412; PubMed=6617663;
RA Baudys M., Kostka V.;
RT "Covalent structure of chicken pepsinogen.";
RL Eur. J. Biochem. 136:89-99(1983).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY, ALTHOUGH BONDS
CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC ALSO CLEAVED TO SOME EXTENT.
CC -1- CATALYTIC ACTIVITY: PREPERENTIAL CLEAVAGE PHE-, LEU-.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
DR PIR: A00984; PECH.
DR HSSP: P00794; 4CMS.
DR MEROPS: A01.001; -;
DR InterPro: IPR001969; Asp_protease.
DR InterPro: IPR001461; pepsin.
DR Pfam: PF00026; asp.1.

DR PRINTS: PRO0792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE; 2.
 KW Hydrolyase; Aspartyl protease; Digestion; Stomach; Zymogen;
 KW Glycoprotein; Gastric juice.
 FT PROPEP 1 367 ACTIVATION PEPTIDE.
 FT CHAIN 43 367 PEPsin A.
 FT ACT_SITE 77 77
 FT ACT_SITE 260 260
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .).
 FT DISULFID 90 95
 FT DISULFID 251 255
 FT DISULFID 290 323
 FT DISULFID 367 367
 FT DISULFID 404 31 MM; 0C547E7FD8F5B341 CRC64;
 SQ SEQUENCE 367 AA; 40431 MM; 0C547E7FD8F5B341 CRC64;
 Query Match 12.5%; Score 302; DB 1; Length 367;
 Best Local Similarity 24.0%; Pred. No. 1.3e-17;
 Matches 88; Conservative 69; Mismatches 125; Indels 84; Gaps 13;
 QY 30 YVEMTVGSPQTLNLTVDGSSNFAVGAAPFPL-----HRYQKQLSTYRDLKRG 82
 DB 59 YGFTISIGTPQDDFSYIEPTGSSNLMV---PSIYKSSACSNHKKRDPKRSSTVSTNET 115
 QY 83 YVVPITQGWBEELGTDIVSIPHPNVTVRANIAITESDK-FETNGSWEGILGLAYAE 141
 DB 116 VYIAIGTSMGSLIGDIYAV---SDIVQNOIFGLSETPGSFYFCNFDILGLAFPS 172
 QY 142 IARPDSLPFFDSLVKQTHV-PNFSIQLCAGFPLNSEVLAVSGSMITGIDHSILY 200
 DB 173 IS--SSGATPVDNMMSQHLVADLFVYLSKDG-----ENGSGVLEGGIDPNYT 220
 QY 201 TGSIMWTPFRREYXEVIIIVREINGODLK--MDCKEYNYKSIYVSGTNTLRLPKVE 258
 DB 221 TGIYVWPLSAETIWOITMDRYVKGKYYACFTC-----QAIYVTGSLVMPGAIN 274
 QY 259 AAVKSIAKASTE-----KPPDGFMLGEQIYWCQAGTTPMNIFFVISLYLMEVTNOS 311
 DB 275 RIKIDGVSDGEISCDISKLPD-----VYFHINGHA----- 307
 QY 312 FRITILPQOTLRVEVDVATISQDCYFAISQSSTG-----VMGAVIMEGFVVDRAR 365
 DB 308 -----FTLPASAYVINEDESCMLGFENMGTPELGEOWILGIVFREYVITDRAN 358
 QY 366 KRIGFA 371
 DB 359 NKVGLS 364
 RESULT 10
 CATE_HUMAN STANDARD; PRT; 396 AA.
 ID CATE_HUMAN STANDARD; PRT; 396 AA.
 AC P14091;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Cathepsin E precursor (EC 3.4.23.34).
 GN CMSE
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89380302; PubMed-2674141;
 RA Akuma T., Pals G., Mohandas T.K., Couvreur J.M., Taggart R.T.;
 RT Human gastric cathepsin E. Predicted sequence, localization to
 RT chromosome 1, and sequence homology with other aspartic
 RT proteinases.;
 RL J. Biol. Chem. 264:16748-16753(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92112877; PubMed-1370478;
 RA Azum T., Liu W.G., Vander Jaan D.J., Bowcock A.M., Taggart R.T.;

RT "Human gastric cathepsin E gene. Multiple transcripts result from
 RT alternative polyadenylation of the primary transcripts of a single
 RT gene locus at 1q31-q32.";
 RL J. Biol. Chem. 267:1609-1614(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Taitnell P.J., Kay J.;
 RL "Human procathepsin E";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 54-68; 77-95; 141-155; 275-285 AND 369-396.
 RX MEDLINE-90241267; PubMed-2334440;
 RA Athauda S.B.P., Matsuzaki O., Kageyama T., Takahashi K.;
 RT "Structural evidence for two isozymic forms and the carbohydrate
 RT attachment site of human gastric cathepsin E.";
 RT Biochem. Biophys. Res. Commun. 168:878-885(1990).
 CC -1 FUNCTION: DUE OF ITS INTRACELLULAR LOCATION AND DISTRIBUTION IN
 CC LYSOPOID ASSOCIATED TISSUE, IT MAY HAVE A ROLE IN IMMUNE FUNCTION.
 CC -1 CATALYTIC ACTIVITY: Similar to cathepsin D, but slightly broader
 CC specificity.
 CC -1 SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M84424; AAA52300.1; JOINED.
 DR EMBL; M84413; AAA52300.1; JOINED.
 DR EMBL; M84417; AAA52300.1; JOINED.
 DR EMBL; M84418; AAA52300.1; JOINED.
 DR EMBL; M84419; AAA52300.1; JOINED.
 DR EMBL; M84420; AAA52300.1; JOINED.
 DR EMBL; M84421; AAA52300.1; JOINED.
 DR EMBL; M84422; AAA52300.1; JOINED.
 DR EMBL; J05036; AAA52300.1; JOINED.
 DR EMBL; A7250717; CAB82850.1; -.
 DR PIR; A34401; A34401.
 DR PIR; A34643; A34643.
 DR PIR; A42038; A42038.
 DR HSSP; P00794; 4CMS.
 DR MEMOPS; A01.010; -.
 DR MIM; 116890; -.
 DR InterPro; IPR001969; Asp-protease.
 DR InterPro; IPR001461; Pepsin.
 DR Pfam; PF00026; asp. 1.
 DR PROSITE; PRO0792; PEPsin.
 DR PRINTS: PRO0792; PEPsin.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolyase; Aspartyl protease; Glycoprotein; Zymogen; Signal.
 FT SIGNAL 1 17
 FT PROPEP 18 53 ACTIVATION PEPTIDE.
 FT CHAIN 54 396
 FT MOD_RES 18 18 PYRROLIDONE CARBOXYLIC ACID.
 FT ACT_SITE 96 96 BY SIMILARITY.
 FT ACT_SITE 281 281 BY SIMILARITY.
 FT DISULFID 60 60 INTERCHAIN (PROBABLE).
 FT DISULFID 109 114 BY SIMILARITY.
 FT DISULFID 272 276 BY SIMILARITY.
 FT DISULFID 314 351 BY SIMILARITY.
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 220 220 O-LINKED (POTENTIAL).
 FT CARBOHYD 333 333 O-LINKED (POTENTIAL).
 SQ SEQUENCE 396 AA; 42793 MM; 40B643C5FB01521E CRC64;
 Query Match 12.4%; Score 299.5; DB 1; Length 396;
 Best Local Similarity 25.9%; Pred. No. 2.3e-17;
 Matches 100; Conservative 65; Mismatches 148; Indels 73; Gaps 16;

EUKARYOTIC ASPARTYL PROTEASES FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: M59235; AA85369.1; -
 DR PIR: C38302; C38302.
 DR HSSP: P00790; IPEN.
 DR MEROPS: A01.001; -
 DR InterPro: IPR001969; Asp_protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp; 1.
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE; 2.
 KW Hydrolyase; Aspartyl protease; Digestion; Stomach; Gastric juice;
 KW Zymogen; Signal; Phosphorylation; Multigene family.
 FT SIGNAL 1 15
 FT PROPEP 16 59
 FT CHAIN 60 387
 FT MOD_RES 129 129
 FT ACT_SITE 93 93
 FT ACT_SITE 276 276
 FT DISULFID 106 111
 FT DISULFID 267 271
 FT DISULFID 310 343
 SQ SEQUENCE 387 AA; 42100 MW; 66FC331A3DC75891 CRC64;

Query Match 12.2%; Score 296; DB 1; Length 387;
 Best Local Similarity 26.9%; Pred. No. 4.3e-17;
 Matches 97; Conservative 63; Mismatches 135; Indels 66; Gaps 13;

QY 30 YVEMTVGSPQPTLNLIVDTGSSNFAVGAAPHF-----LHRYQROLSTYRLKRG 82
 DB 75 YGCTISIGRPQDFYVIFDTGSSNLMV---PSTYCSLACALHKRNPEDSSRYQCTSEI 131
 QY 83 VVPTVQGWESLGTDLVSIPIGPNVTVRANIAITESDKPFINSNMGGILGLAYAEI 142
 DB 132 LSTVGTGSGMTGLGYDVKVGSIEDTNOIFGLSKTEPSLTFLE--APDGLIGLAVPSI 189
 QY 143 ARPDLSLEFPFSLVKQFHV--PULFSLQCGAFPLNQSSEVLASVGSIMIGIDSLVLT 201
 DB 190 SSSDAI--PVFDMMNMGILVSODLFSVYSSDD-----EKSLVFGIDISYTT 237
 QY 202 GSIWYPIRREWEYEVIVRVEINGDLKM--DCKEYNTDKSIVDSGTNLRLLPKVFPEA 259
 DB 238 GSIWVPAVSEGVQWQITMDSVINGETIACADSC-----QAIVDTGSLTGP---TS 287
 QY 260 AKSITAASTEFPPGFLVGLGELY--CMQAGTTPNNIIFPISILYLMGEVYNGSFRITLIP 318
 DB 288 AISNIQSYIGASR---NLIGENVISCSAIDSLDPIV-----TING 325
 QY 319 QOYLRFVEVDVATSDQCYFAISOSSTGT-----VMGAVIMEGFVEVDEPRARRKIGFAV 372
 DB 326 IQYPLPASAVIILKEDDCTSGLEGMMVDTYTGELMILGIVFINQYFTVDPDRANNGGLAA 385
 QY 373 S 373
 DB 386 A 386

RESULT 13
 CATE_CAVPO STANDARD; PRT; 391 AA.
 ID CATE_CAVPO
 AC P25796;
 DT 01-MAY-1992 (rel. 22, Created)
 DT 01-MAY-1992 (rel. 22, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)

DE Cathepsin E precursor (EC 3.4.23.34).
 GN CTSB.

OS Cavia porcellus (Guinea pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Hystriognathii; Cavidae; Cavia.
 OX NCBI_Taxid=10141;

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RA MEDLINE=92355614; PubMed=1644829;
 RX Kageyama T., Ichinose M., Tsukada S., Miki K., Kurokawa K., Koiwai O.,
 RA Tanji M., Yakabe E., Athauda S.B., Takahashi K.;
 RT "Gastric procathepsin E and progastricsin from guinea pig.
 RT Purification, molecular cloning of cDNAs, and characterization of
 RT enzymatic properties, with special reference to procathepsin E.";
 RL J. Biol. Chem. 267:16450-16459(1992).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Gastric mucosa;
 RX MEDLINE=96073637; PubMed=8540321;
 RA Kageyama T., Ichinose M., Miki K., Moriyama A., Yonezawa S.,
 RA Tanji M., Athauda S.B., Takahashi K.;
 RT "Isolation, characterization, and structure of procathepsin E and
 RT cathepsin E from the gastric mucosa of guinea pig.";
 RL Adv. Exp. Med. Biol. 362:211-221(1995).

CC -! FUNCTION: DUE TO ITS INTRACELLULAR LOCATION AND DISTRIBUTION IN
 CC LYMPHOID ASSOCIATED TISSUE, IT MAY HAVE A ROLE IN IMMUNE FUNCTION.
 CC -! CATALYTIC ACTIVITY: Similar to cathepsin D, but slightly broader
 CC specificity.

CC -! SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: M88653; AA37052.1; -
 DR EMBL: S80547; AA35844.1; -
 DR PIR: A43356; A43356.
 DR HSSP: P00794; 4CMS.
 DR MEROPS: A01.010; -
 DR InterPro: IPR001969; Asp_protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp; 1.
 DR PRINTS: PR00792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE; 2.
 KW Hydrolyase; Aspartyl protease; Glycoprotein; Zymogen; Signal.
 FT SIGNAL 1 18
 FT PROPEP 19 53
 FT CHAIN 54 391
 FT ACT_SITE 92 92
 FT ACT_SITE 276 276
 FT DISULFID 56 56
 FT DISULFID 105 110
 FT DISULFID 267 271
 FT DISULFID 309 346
 FT CARBOHYD 86 86
 SQ SEQUENCE 391 AA; 42132 MW; 78D216BF8CFDABD CRC64;

Query Match 12.2%; Score 296; DB 1; Length 391;
 Best Local Similarity 26.9%; Pred. No. 4.4e-17;
 Matches 98; Conservative 64; Mismatches 130; Indels 72; Gaps 16;

QY 30 YVEMTVGSPQPTLNLIVDTGSSNFAVGA---APHPFLHRYQROLSTYRLKRGVY 85
 DB 74 YGCTISIGSPQNFYVIFDTGSSNLMVPSVYCTSPACQTHPVHPISLSYREVGNFSI 133
 QY 86 PYTQGWESLGTDLVSIPIGPNVTVRANIAITESDKPFINSNMGGILGLAYAEIAR 145

DB 134 QYNGSLTGTIGADQVSV-EGLTIVGQGFESVQEPKTFVH-AEFGILGLGYPDLAA-190
QY 146 DDLSEPFEDLVQTHPNLFFSLQLCAGFPPLNQSEVLAVGGSMITGIDHSLYGSLSW 205
DB 191 -GGVTFPFDMMKMQ-----NVALPML-----FSVYSSNPGSGSSELFFGGYDPSHFGSLN 241
QY 206 YTEIRREMYEVLIVVEINGODLKMCKEYNDKSIYDSGTINLRLPKVFEEAAKSTK 265
DB 242 WVPVTKQAVWQIALDGIQVQ--DSVFCSP--GCQAVIDGTSLTGP-----FGKITQIQ 293
QY 266 AASTKEFPDGFELGEOLYCWAQATTPWNIPIVLSLYMGEVTSNOSFRI-----TILPQ 320
DB 294 EALGATYVDEGY-----SVQC-----ANLNMMLDVT-----FLINGVPTLNP 333
QY 321 YLRPEDEVATSGDDCYKFAISOSTG-----TYMAYIMEGFYVVEDARRK 367
DB 334 Y--TLIDFVDGMQVC-----STGFEGLEIQPPAGPLMILGDVPIROFYAVFDRGNMR 383
QY 368 IGFA 371
DB 384 VGIA 387

RESULT 14
CATD_HUMAN
ID CATD_HUMAN STANDARD; PRT; 412 AA.
AC P07339;
DT 01-APR-1988 (rel. 07, Created)
DT 01-APR-1988 (rel. 07, Last sequence update)
DT 01-MAR-2002 (rel. 41, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
GN CTSD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NX NCBI_Taxid=9606;
RY [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270436; PubMed=3927292;
RA Faust P.L., Kornfeld S., Chirgwin J.M.;
RT "Cloning and sequence analysis of cDNA for human cathepsin D.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4910-4914(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87231068; PubMed=3588310;
RA Westley B.R., May F.E.B.;
RT "Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive
RL human breast cancer cells.";
RN Nucleic Acids Res. 15:3773-3786(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91299158; PubMed=2069717;
RA Redeker B., Heckendorf B., Grosch H.W., Mersmann G., Hasilik A.;
RT "Molecular organization of the human cathepsin D gene.";
RL DNA Cell Biol. 10:423-431(1991).
RN [4]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE=94085791; PubMed=8262386;
RA May F.E., Smith D.J., Westley B.R.;
RT "The human cathepsin D-encoding gene is transcribed from an estrogen-
RL regulated and a constitutive start point.";
RN Gene 134:277-282(1993).
RN [5]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE=95021301; PubMed=7935485;
RA Augereau P., Mitrallas F., Cavailles V., Gaudelot C., Parker M.,
RA Rochofort H.;
RT "Characterization of the proximal estrogen-responsive element of
RL human cathepsin D gene.";
RN Mol. Endocrinol. 8:693-703(1994).
RN [6]
RP SEQUENCE OF 170-180.

RC TISSUE=Liver;
RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
RA Appel R.D., Hughes G.J.;
RL Submitted (JUN-1992) to the SWISS-PROT data bank.
RN [7]
RP VARIANT VAL-58.
RX MEDLINE=20179010; PubMed=10716266;
RA Pappasoliotopoulos A., Bagli M., Kurz A., Kornhuber J., Forstl H.,
RA Maier W., Pauls J., Lautenschlager N., Heun R.;
RT "A genetic variation of cathepsin D is a major risk factor for
RL Alzheimer's disease.";
RN Ann. Neurol. 47:399-403(2000).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RC TISSUE=Spleen;
RX MEDLINE=93223670; PubMed=8467789;
RA Metcalf P., Fusek M.;
RT "Two crystal structures for cathepsin D: the lysosomal targeting
RL signal and active site.";
RN EMBO J. 12:1293-1302(1993).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93342076; PubMed=8393577;
RA Baldwin E.T., Bhat T.N., Gulink S., Hosur M.V., Sowder R.C. II,
RA Cachau R.E., Collins J., Silva A.M., Erickson J.W.;
RT "Crystal structures of native and inhibited forms of human cathepsin
RL D: implications for lysosomal targeting and drug design.";
RN Proc. Natl. Acad. Sci. U.S.A. 90:6796-6800(1993).
CC -1- FUNCTION: CATHEPSIN D IS AN ACID PROTEASE ACTIVE IN INTRACELLULAR
CC PROTEIN BREAKDOWN.
CC -1- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC that of pepsin A. Does not cleave the 4-Gln-1-His-5 bond in B
CC chain of insulin.
CC -1- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC -1- SUBCELLULAR LOCATION: Lysosomal.
CC -1- POLYMORPHISM: The Val-58 allele is significantly overrepresented
CC in demented patients (11.8%) compared with nondemented controls
CC (4.9%). Carriers of the Val-58 allele have a 3.1-fold increased
CC risk for developing AD than noncarriers.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11233; AAB59529.1; -;
DR EMBL; X05344; CAA28955.1; -;
DR EMBL; M63138; AAA51922.1; -;
DR EMBL; M63134; AAA51922.1; JOINED.
DR EMBL; M63135; AAA51922.1; JOINED.
DR EMBL; M63136; AAA51922.1; JOINED.
DR EMBL; M63137; AAA51922.1; JOINED.
DR EMBL; L12980; AAA16314.1; -;
DR EMBL; S74689; AAD14156.1; -;
DR EMBL; S52557; AAD13868.1; -;
DR PIR; A25771; KHHUD.
DR PDB; 1LYA; 31-JAN-94.
DR PDB; 1LYB; 31-JAN-94.
DR MEROPS; A01.009; -;
DR SWISS-2DPAGE; P07339; HUMAN.
DR SILEN-2DPAGE; P07339; -;
DR MTM; 116840; -;
DR InterPro; IPR001969; Asp-Protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 1.
DR PRINTS; PR00792; PEPSIN.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2002, 09:13:36 ; Search time 32.31 Seconds

(without alignments)
1356.135 Million cell updates/sec

Title: US-09-724-571-43

Perfect score: 2419
Sequence: 1 ETDEEPEEPGRGSRGVEMVD.....CLRCLRQHQHDFADISILK 456

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2419	100.0	501	2	A59090 aspartic proteinase
2	308	12.7	384	2	JC7574 pepsinogen A - Afr
3	308	12.7	385	2	JC7575 pepsinogen A - bul
4	306	12.6	387	2	B38302 pepsin (EC 3.4.23.
5	303	12.5	383	2	JC7573 pepsinogen C - Afr
6	302.5	12.5	388	1	S19682 pepsin A (EC 3.4.2
7	301	12.4	382	1	PECH pepsin A (EC 3.4.2
8	299.5	12.4	396	2	A34401 cathepsin E (EC 3.
9	298.5	12.3	383	2	A41443 pepsin (EC 3.4.23.
10	298.5	12.3	384	2	A39314 gastricsin (EC 3.4
11	296	12.2	387	2	C38302 pepsin (EC 3.4.23.
12	296	12.2	391	2	A43356 cathepsin E (EC 3.
13	295.5	12.2	412	1	KHHUD pepsin D (EC 3.
14	295	12.2	387	2	D38302 pepsin (EC 3.4.23.
15	289.5	12.0	388	1	S19684 pepsin A (EC 3.4.2
16	287.5	11.9	444	2	T24204 hypothetical prote
17	287	11.9	407	1	KHRPD cathepsin D (EC 3.
18	286	11.8	387	2	E38302 pepsin (EC 3.4.23.
19	285.5	11.8	398	2	S66455 cathepsin E (EC 3.
20	285	11.8	398	2	I51185 cathepsin D (EC 3.
21	283.5	11.7	388	1	PEMOCAR pepsin A (EC 3.4.2
22	282.5	11.7	388	1	PEHU pepsin A (EC 3.4.2
23	281.5	11.6	381	1	CMSHR pepsin A (EC 3.4.2
24	281.5	11.6	388	1	PEMOCJ pepsin A (EC 3.4.2
25	281.5	11.6	388	1	A50142 pepsin A (EC 3.4.2
26	279.5	11.6	388	2	B30142 pepsin A (EC 3.4.2
27	279.5	11.6	410	1	KHMSD cathepsin D (EC 3.
28	278.5	11.5	386	1	PEPG pepsin A (EC 3.4.2
29	278	11.5	387	2	JC7245 pepsinogen A - com

30	277.5	11.5	380	2	I47176 chymosin (EC 3.4.2
31	277.5	11.5	396	2	S36865 cathepsin E (EC 3.
32	276	11.4	389	2	JEO371 pepsin C (EC 3.4.2
33	273.5	11.3	381	1	CMBO pepsin C (EC 3.4.2
34	270.5	11.2	377	1	PEMOCJ gastricsin (EC 3.4
35	270.5	11.2	389	2	A38302 pepsin (EC 3.4.23.
36	270	11.2	376	2	I45856 aspartic proteinase
37	268.5	11.1	344	1	KHPGD cathepsin D (EC 3.
38	267.5	11.1	381	2	JC7247 prochymosin - comm
39	266	11.0	380	2	S03433 candidapepsin (EC
40	266	11.0	405	2	A25379 saccharopepsin (EC
41	264	10.9	396	2	T47207 aspartic proteinase
42	263.5	10.9	388	2	JC7246 pepsinogen C - com
43	262.5	10.9	394	2	B43356 gastricsin (EC 3.4
44	261.5	10.8	387	2	A45117 aspartic proteinase
45	261.5	10.8	388	2	A29937 gastricsin (EC 3.4

ALIGNMENTS

RESULT 1

A59090 aspartic proteinase (EC 3.4.23.-) BACE precursor - human

N:Alternate names: beta-secretase; beta-site APP cleaving enzyme
C:Species: Homo sapiens (man)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-May-2000

C:Accession: A59090
R:Vassar, R.; Bennett, B.D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E.A.; Denis, P.; Teplio

M.A.; Biere, A.L.; Curran, E.; Burgess, T.; Louis, J.C.; Collins, F.; Treanor, J.; Ro
S:Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the tran
A:Reference number: A59090; MUID:20002972
A>Note: submitted to Genbank, September 1999

A:Accession: A59090
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-501 <VAS>

A:Cross-references: GB:AF190725; NID:G6118538; PID:NARF04142.1; PID:G6118539
C:Genetics:

A:Gene: BACE
C:Superfamily: beta-secretase
C:Keywords: Alzheimer's disease; aspartic proteinase; brain; glycoprotein; hydrolase;
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-45/Domain: propeptide #status predicted <PRO>
F:46-501/Product: acid proteinase BACE #status predicted <MAT>
F:461-477/Domain: transmembrane #status predicted <TRN>

F:93,289/Active site: Asp #status predicted
F:153,172,223,354/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:330-380/Disulfide bonds: #status predicted

Query Match 100.0% Score 2419; DB 2; Length 501;
Best Local Similarity 100.0% Pred. No. 3.9e-192;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ETDEEPEEPGRGSRGVEMVDNLGRSGGGYEMVGSPPOTLNTLVYDSSNFVAGAP	60
DB	46	ETDEEPEEPGRGSRGVEMVDNLGRSGGGYEMVGSPPOTLNTLVYDSSNFVAGAP	105
QY	61	HPFLHRYQROLSTYRDLRKGVYVYTGKWEGLGTDLVSIHPGPNVYVRANITATTE	120
DB	106	HPFLHRYQROLSTYRDLRKGVYVYTGKWEGLGTDLVSIHPGPNVYVRANITATTE	165
QY	121	SDKEPFGSNNEGILGLAVAEIARDDSLPEPDSLVQTHVPLNFSIQLCGAGPPLNOS	180
DB	166	SDKEPFGSNNEGILGLAVAEIARDDSLPEPDSLVQTHVPLNFSIQLCGAGPPLNOS	225
QY	181	EVLASVGGSMIIIGGIDHSLYTGSLWYTPIRREMYEVIIVREINIGODLKMDCREYNDK	240
DB	226	EVLASVGGSMIIIGGIDHSLYTGSLWYTPIRREMYEVIIVREINIGODLKMDCREYNDK	285
QY	241	SIVDSGTTNLRPLPKKVFPAVAKSIKASSTEFDPDGFWIGQLVCMQAGTTPWNIIPVIS	300

```

|||||
Db 286 STVDSGTNLRPKVFFRAVAKSSTKFAASSTKEFPDGFHLEGLVCMQAGTTPWNIFFPVIS 345
QY 301 LYLMGEVTSNOSFRITLIPQOYLIRPEVDVATSODDCYKFAISOSSTGYVMGAVIMEGFYV 360
Db 346 LYLMGEVTSNOSFRITLIPQOYLIRPEVDVATSODDCYKFAISOSSTGYVMGAVIMEGFYV 405
QY 361 FDRARRRIGFAVSACHVDEFTFAVEGPFVTLDMEDCGYNIPQDESTLMTIAYVMAI 420
Db 406 FDRARRRIGFAVSACHVDEFTFAVEGPFVTLDMEDCGYNIPQDESTLMTIAYVMAI 465
QY 421 CALFMPLPLCMWCQWRCLRCLRQHDHDFADISLKL 456
Db 466 CALFMPLPLCMWCQWRCLRCLRQHDHDFADISLKL 501

RESULT 2
JC7574
pepsinogen A - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C:Accession: JC7574; PC7119
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens
A:Reference number: JC7573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: JC7574
A:Molecule type: mRNA
A:Residues: 1-384 <IKU>
A:Cross-references: DDBJ:AB045380
A:Accession: PC7119
A:Molecule type: protein
A:Residues: 16-35,57-76 <IK2>
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like a
C:Genetics:
A:Gene: Pga
C:Superfamily: pepsin
C:Keywords: stomach; zymogen

Query Match 12.7%; Score 308; DB 2; Length 384;
Best Local Similarity 27.5%; Pred. No. 1.3e-17;
Matches 99; Conservative 60; Mismatches 133; Indels 68; Gaps 16;

Db 72 YVETISIGTPPQEFYVIFDGTGSANLWV--PSVYSSQACSNNRNPQSSSTFQATNTP 128
QY 30 YVEMTVGSPPTQTLNLTVDTGSSNFAVGAAPHPFL-----HRYQRQLSTYRDLRKG 82
Db 72 YVETISIGTPPQEFYVIFDGTGSANLWV--PSVYSSQACSNNRNPQSSSTFQATNTP 128
QY 83 VVVPYQGWEGELGTDLVSIPHGPNVTVRANIAITESDK-FFINGSNMGEILGLAVAE 141
Db 129 VSIQYOTGSMSEFLGVDITQV---GNIQISNQMFGLSEBPSFLYSPFDGILGLAFPS 185
QY 142 IARPDDSLPEPFDLVKQTHVP-NLFSQLCGAGPPLNQSEVLASVSGSMIIGIDHSLY 200
Db 186 IA--SSQATPVPDNNMSQGLIPQNLFSVYLSPG-----QTSYVLEFGGVNLSY 233
QY 201 TCSLWATPIRREMYEVIIIVREINGQDL--KMDCKEYNDKSIYVSGTTLRLPKVEE 258
Db 234 SGLSNVPLTAETIYQITLDSVINSQVACSSC-----QAIYTFGSLMGPSTPI- 286
QY 259 AAVKSIRKASSTKEFPDGFHLEGLVCMQAGTTPWNIFFPVISLYLMGEVTSNOSFRITL 317
Db 287 ANIQYIGASQSSN-----GQYVINCNNISMPTIVF-----TTN 321
QY 318 POQY-LRPEDVATSODDCYK-FAISOSSTGT-----VMGAVIMEGFYVDFARRRIGFA 371
Db 322 GVOYPLSPAYVRQNOGSSGFGQANMLPTNSGDLWILGDFIRYQYTFVDFRANNTYVAIA 381

RESULT 3
JC7575
pepsinogen A - bullfrog

```

```

C:Species: Rana catesbeiana (bullfrog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C:Accession: JC7575
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinog
A:Reference number: JC7573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: JC7575
A:Molecule type: mRNA
A:Residues: 1-385 <IKU>
A:Cross-references: DDBJ:AB045376
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-lik
C:Genetics:
A:Gene: Pga
C:Superfamily: pepsin
C:Keywords: stomach; zymogen

```

```

Query Match 12.7%; Score 308; DB 2; Length 385;
Best Local Similarity 27.8%; Pred. No. 1.3e-17;
Matches 99; Conservative 64; Mismatches 133; Indels 60; Gaps 15;

QY 30 YVEMTVGSPPTQTLNLTVDTGSSNFAVG---AAPHPFLHRYQRQLSSTYRDLRKGYYV 85
Db 73 YVETISIGTPPQEFYVIFDGTGSANLWVPSVYSSQACSTNHHFNNQSSSTFQATNPVSI 132
QY 86 PYTQGWEGELGTDLVSIPHGPNVTVRANIAITESDK-FFINGSNMGEILGLAVAEIAR 144
Db 133 QYGTGSMSEFLGVDITQV---GNIQITNOIFGLSSEPSFLYSPFDGILGLAFPSLA- 188
QY 145 PDDSLPEPFDLVKQTHVP-NLFSQLCGAGPPLNQSEVLASVSGSMIIGIDHSLYGS 203
Db 189 -SSQATPVPDNNMSQGLIPQDLFSVYLSQG---QS-----GSFVLEFGVDITYTGN 237
QY 204 LMYTPIRREMYEVIIIVREINGQDLK--DCKEYNDKSIYDSTGTLRLPKVEEAYV 261
Db 238 LNWVPLTAETIYQITVDSISIGQVYACSGSC-----SAIVDTSTSLAGP-----STPI 287
QY 262 KSIKASSTKEFPDGFHLEGLVCMQAGTTPWNIFFPVISLYLMGEVTSNOSFRITLIPQY 321
Db 288 ANIQYIGANQSSNGQYV---INCNNISMPTIVF-----TINGVQY 326
QY 322 LRPVED-VATSODDC--YKFAISOSSTGT--VMGAVIMEGFYVDFARRRIGFA 371
Db 327 PLPASAYVRQSOQSCSTSGFGQANMLPTNSGDLWILGDFIRYQYTFVDFRANNTYVAMA 382

RESULT 4
B38302
pepsin (EC 3.4.23.-) II-1 precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Jun-1991 #sequence_revision 20-Sep-1991 #text_change 23-Feb-1997
C:Accession: B38302
R:Kaageyama, T.; Tanabe, K.; Koizumi, O.
J. Biol. Chem. 265, 17031-17038, 1990
A:Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nu
A:Reference number: A38302; MUID:91009127
A:Accession: B38302
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-387 <KAG>
A:Cross-references: GB:M59235; GB:J05638
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion

Query Match 12.6%; Score 306; DB 2; Length 387;
Best Local Similarity 27.1%; Pred. No. 2e-17;
Matches 98; Conservative 67; Mismatches 131; Indels 66; Gaps 15;

QY 30 YVEMTVGSPPTQTLNLTVDTGSSNFAVG---AAPHPFLHRYQRQLSSTYRDLRKGYYV 85

```

D 75 YFCTISGTPPEFTVIEPTGSSNLNVPSTYSSLACFLHKRPDDSTFQATSETLSI 134
Q 86 PYQGMKEGELGTDLVSIPIHPNVTVRANIATTESD---KEPINSNMEGILGLAYAEI 142
D 135 YVGTGSMGTGLGIDYKVV---GNIEDTNGIIFGSKTEPGITFLV--APDGLIGLAPSEI 189
Q 143 ARPDDELPEFPDVLVQTHV-PNLFSLQICGAGFPLNOSEVLASVGSMTIGIDHSLYT 201
D 190 SASDAT--PVFDNMNMEGVSEDLFVSYLSSNG-----EKSMVFGIDSSYTT 237
Q 202 GSLMYPIRREMYEVIYIRVEINSGDLK--DCKEYNKDSIVDSGTNLRPKVFEFA 259
D 238 GSLMWVPSVHEGWQITMSTITINGETIACADSC-----QAVVDGTSLLAGPISAIK 291
Q 260 AVKSIKAASSTEFKPDGFWLGEOLV-CWQAGTTPMNIFFVILSYLMGEVTNOSFRITILP 318
D 292 IQSYIGASKNL-----LGENITCSAIDSLPDIVF-----TINN 325
Q 319 QOYLRFVED-VATSDDC---TKFAISOSTGT--VMGAVIMEGFYVFDRAKRIGFAV 372
D 326 VQYPLPASAVILKEDDCLSGFDGMNLDTSYGEMLTGDFVTFROYFTVDRANNOVGLAA 385
Q 373 SA 374
D 386 AA 387

RESULT 5
JC7573
pepsinogen C - African clawed frog
N:Alternate names: progastrolisin
C:Species: xenopus laevis (African clawed frog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C:Accession: JC7573; PC7118
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens
A:Reference number: JC7573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: JC7573
A:Molecule type: mRNA
A:Residues: 1-383 <IKU>
A:Cross-references: DDBJ:AB045379
A:Accession: PC7118
A:Molecule type: protein
A:Residues: 17-68 <IK>
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like a
C:Genetics:
A:Gene: PGC
C:Superfamily: pepsin
C:Keywords: stomach; zymogen

Query Match 12.5%; Score 303; DB 2; Length 383;
Best Local Similarity 27.4%; Pred. No. 3.4e-17;
Matches 104; Conservative 57; Mismatches 122; Indels 96; Gaps 17;
Q 30 YVENVTSPPQTLNLIYDVGSSNFANVA-----APHFLHRYXORLSSTYRDLRK 81
D 67 YVGEISITGPONLVLFDVTGSSNLNVAASYCOSQACTNHL---FNPSSSTYSSNQ 122
Q 82 GVYVYVYTGKKEGELGTLVSIPIHPNVTVRANIATIESKFTINGSN---WEGTLG 136
D 123 QFSLQIGYGSGLTGLIGYTVTI---QNVAISQOEGFSETP---GTNEVYAOFDGLG 175
Q 137 LAYEIAIAPDLSLEFPDVLVQTHV-PNLFSLQICGAGFPLNOSEVLASVGSMTIGID 196
D 176 LAYPSIA--VGGATTVMQGMQO---NLNQPI--FGFYLSGS--SQNGEVAIFGVD 225
Q 197 HSLTGSIMYPIRREMYEVIYIRVEINSGD---LKMCKEYVYDKSIVDSGTNLRPL 253
D 226 ONYTTGQIYVTPVSEYVWQIGQFSGINGATGWCSCGC-----QATVDGTGTSLLTAP 279

Q 254 KRYEAAVAKSIKAASSTEFKPDGFWLGEOLVCMQAGTTPMNI--FPVILSYLMG----- 305
D 280 QVFSFSLISDIAQODQN-----GQIVVSCS-----NQNUPITISFTISGVSFPLP 325
Q 306 ---EVTNOS-----FRITLPOQYLRFVEDVATSDCKYFAISOSTGTVMGAVIME 355
D 326 PAAVYVLOQSSGYCTIGIMPTVLPSONGQPL-----WILGDVFLR 364
Q 356 GFYVVEDRAKRIGFAVSA 374
D 365 EYYSVYDLGNNOVGPAATA 383

RESULT 6
S19682
pepsin A (EC 3.4.23.1) 4 precursor - Japanese macaque
N:Alternate names: pepsinogen A isozyme 4
C:Species: Macaca fuscata (Japanese macaque)
C:Date: 22-Nov-1993 #sequence_revision 19-Oct-1995 #text_change 18-Jun-1999
C:Accession: S19682; S16065
R:Kageyama, T.; Tanabe, K.; Koike, O.
Eur. J. Biochem. 202, 205-215, 1991
A:Title: Development-dependent expression of isozymogens of monkey pepsinogens and st
A:Reference number: S19681; MUID:92037645
A:Accession: S19682
A:Molecule type: mRNA
A:Residues: 1-388 <KAG>
A:Cross-references: EMBL:X59753; NID:938070; PIDN:CAA42425.1; PID:938071
A:Note: parts of sequence, including amino ends of pepsinogen and activation intermedi
C:Comment: This is a minor component of pepsin at all post-partum stages.
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; gastric juice; hydrolyase; phosphoprotein; protein di
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-388/Product: pepsinogen A 4 #status experimental <PPR>
F:16-62/Domain: activation peptide #status experimental <APR>
F:63-388/Product: pepsin A 4 #status experimental <ENZ>
F:38-39/Cleavage site: Leu-Lys (pepsin) #status experimental
F:62-63/Cleavage site: Leu-Ile (pepsin) #status experimental
F:94,277/Active site: Asp #status predicted
F:107-112,268-272,311-344/Disulfide bonds: #status predicted
F:130/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 12.5%; Score 302.5; DB 1; Length 388;
Best Local Similarity 27.6%; Pred. No. 3.8e-17;
Matches 107; Conservative 64; Mismatches 133; Indels 83; Gaps 17;
Q 3 DEEPEPGRGSEVEMVDNLNKGSGQGYVEMTVGSPPTLNLIVDTGSSNFAVGAAPHP 62
D 64 DEQPLE-----NYLDV-----EYFGTIGTGPQANTVTFEDGSSNLWV---PSV 105
Q 63 FL-----HRYXORLSSTYRDLRKGYVYPIYTGKKEGELGTDLVSIPIHPNVTVRANI 115
D 106 YCYSLACMDHNLFPDOSTYRATSKVTSITYGSGMTGLIGYTVKVV---GGISDTPNQI 162
Q 116 AAITESDK-FEINGSNMEGILGLAYAEIARPDLSLEFPDVLVQTHV-PNLFSLQICGA 173
D 163 FGLSETEPGFFLYIARPDGLIGLAPYSI--SSGATVFDIMNQRLVSDLSVYSIAD 220
Q 174 GFPLNOSEVLASVGSMTIGIDHSLTGSIMYPIRREMYEVIYIRVEINSGDL--KM 231
D 221 ---DOS-----GSVYIFGIDSSYTTGSLNMPVSEVGEWQISVDSITMNGTICAK 270
Q 232 DCKEYNKDSIVDSGTNLRPKVFEFAAVKSIKAASSTEFKPDGFWLGEOLV-CWQAGT 290
D 271 GC-----QATVDGTSLTGTPSPINNIOSDIGASNSD-----GEMVYVSCSAISS 316
Q 291 TPWNIFFVILSYLMGEVTNOSFRITILPOOY-LRPVEDVATSDDCYK-----FAISOSS 344
D 317 LPDIVF-----TINGVQYPLPSPAYIIQSGSGCTSGRGADVPTESG 358
Q 345 TGYVMGAVIMEGFYVFDRAKRIGFA 371


```

Db      170 TEPOQTETVD-AEFDGIIIGLGYPSLA--VGCVTPVFDNMMAQ---NLVDPMPMSVYSSN 222
QY      179 QSEVLASVGSMTIGDHSIYTSGLWYTPIRRMRYEVLIVREINQGLDKMCKEYNY 238
Db      223 PE---GGAGSELIFGGYDHFSGSLMWVPVTKQAYWQALDNIQYVG--TVMECSF--G 275
QY      239 DKSTVDSGTTNLRLPKRVFEAAVSKIRASSTEFKPDGFMGLGRLVCWQAGTTPMWNFPV 298
Db      276 COAIVDTGTSITGSPDKIKQLQNALCAAP-----VDGEVAVE-----CANLVNMPD 322
QY      299 ISLVLMGEVNVNSFRITILLQOYLKRPVEDVATSDODCYKFAISQSTG----- 346
Db      323 VTFTING-----VPRYLSPTAY--TLDPVDMQFC-----SSGFGGLDIHPAPG 366
QY      347 -TVAGAVIMEGFYVFPDRARRKRIEGA 371
Db      367 LWILGDVFIRQFYSVPRGNRRVGLA 392

```

```

RESULT 9
A41443
pepsin (EC 3.4.23.-) precursor, embryonic - chicken
C:Species: Gallus gallus (chicken)
C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 21-Jul-2000
C:Accession: A41443
R:Hayashi, K.; Agata, K.; Mochii, M.; Yasugi, S.; Eguchi, G.; Mizuno, T.
J. Biochem. 103, 290-296, 1988
A:Title: Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken pepsin
A:Reference number: A41443; MUID:86227903
A:Accession: A41443
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-383 <NAY>
A:Cross-References: GB:D00215; NID:92760810; PIDN:BA00153.1; PID:g222853
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; protein digestion

```

```

Query Match      12.3%; Score 298.5; DB 2; Length 383;
Best Local Similarity 25.2%; Pred. No. 8e-17;
Matches 90; Conservative 75; Mismatches 125; Indels 67; Gaps 14;

QY      30 YVEMTVGSPQTLNLTIVDTGSSNFAVGA---APHPFLARKYQQLSSTYDRLKRYIV 85
Db      76 YGTGISTGTPPDFTYVFTGSSNLMWPSVCTSPACQSHQMFNSQSTYKSTGQNTLSI 135
QY      86 PYTGKMEGELGTDLVSIPIHGPNVTVRANIAITESDKFFINGSMWEGILGLAVALARP 145
Db      136 HGTGDMECTVGCDFVTYVASLMDTNQLFGLST-SFPGQFVY-VKFDGILGLGYSLAA- 192
QY      146 DDLSEFPDLSVKQTHV-PNLFSLDLCGAFPLNQSEVLASVGSMTIGDHSIYTSGL 204
Db      193 -GGITPVEDNMVNESLLEQNLFYSVLS-----REPMGSNVVPGIDESYFTGSI 240
QY      205 WYTPIRREYVYVIVREINQGLD--KMDCKEYVYKSYVDSGTNLRPKRYFEAAVK 262
Db      241 NMIPTSYGYQWQISMDSTIVNQELTACSSG-----QALIDTGSILVAGAPASINDIOS 294
QY      263 SIKAASTEFKPDGFMGLGRLVCWQAGTTPMWNFPVLS---YLMGEVTVNSFRITILLP 318
Db      295 AVGANQNT-----YGEYSV-----NCSHILAMPDVVFYIGI----- 326
QY      319 OQYLKRPVEDVA---TSQDDCKKFAISQSSGTGVAGAVIMEGFYVFPDRARRKRIEGA 371
Db      327 -QY--PVPAALATTEONGGTGTCMSSFQNSSADLWILGDVFIRYYSIFDRANNRVGLA 380

```

```

RESULT 10
A39314
gastricain (EC 3.4.23.3) precursor - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 22-Jun-1999

```

```

C:Accession: A39314
R:Yakabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Miki, K.; Kurokawa, K.; Ito, H.; Kag
J. Biol. Chem. 266, 22436-22443, 1991
A:Title: Purification, characterization, and amino acid sequences of pepsinogens and
A:Reference number: A39314; MUID:92042186
A:Accession: A39314
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <YAK>
A:Cross-References: GB:M73750; NID:9213687; PIDN:AAA49530.1; PID:g213688
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; protein digestion

```

```

Query Match      12.3%; Score 298.5; DB 2; Length 384;
Best Local Similarity 25.5%; Pred. No. 8.1e-17;
Matches 100; Conservative 61; Mismatches 142; Indels 89; Gaps 15;

```

```

QY      14 SFVEMVDNLKSGSGGYVEMTVGSPQTLNLTIVDTGSSNFAVGAAPHPFL-----HR 66
Db      51 NPAIAFEPLANTMDMSYGEISIGTPQNFVLFPDTGSSNLMV--PSTYCQSQACTNHP 107
QY      67 YVQRLSSTYRDLRKGVVYPTQGWKEGELGTDLVSIPIHGPNVTVRANIA-----AI 118
Db      108 QPNPQSSSYSSNQOQFSLQYCTGSLTGLGIDYVQI-----QNALSQQDEGLSV 158
QY      119 TSDKFFINGSMWEGILGLAVALARPDSLEFPDLSVKQTHVPN-LFSLQCGAGFPL 177
Db      159 TEPGTFYV-AQFDGIIIGLAVPSIA--EGGATTVWQMIQNLNLQPLFARYLSGQSQMS 215
QY      178 NOSEVLASVGSMTIGDHSIYTSGLWYTPIRRMRYEVLIVREINQGLD---LKMDCK 234
Db      216 N-----GGEVAVGVDQNTVTSQIYTPVTSFYQWQIQGFSVNGATGMCSSQC- 266
QY      235 EYVNDKSYVDSGTTNLRLPKRVFEAAVSKIRASSTEFKPDGFMGLGRLV-CWQAGTTPW 293
Db      267 -----QGIVDTGTSITGSPDKIKQLQNALCAAP-----VDGEVAVE----- 313
QY      294 NIFPVI-----SLYLMGEVTVNSFRITILLPQOYLKRPVEDVATSDODCYKFAISQ 342
Db      314 ISFTISGVSPPLPSAVYVLAQNSGYCTGTGIMPTYLPSPONGPL----- 356
QY      343 SSTGTWGAVIMEGFYVFPDRARRKRIEFAVSA 374
Db      357 ----WILGDVFLROYYSYVDLGNNOYVGAANA 384

```

```

RESULT 11
C38302
pepsin (EC 3.4.23.-) II-2/3 precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 23-Feb-1997
C:Accession: C38302
R:Kageyama, T.; Tanabe, K.; Koizumi, O.
J. Biol. Chem. 265, 17031-17038, 1990
A:Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nu
A:Reference number: A38302; MUID:91009127
A:Accession: C38302
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-387 <KAG>
A:Cross-References: GB:J05638
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion

```

```

Query Match      12.2%; Score 296; DB 2; Length 387;
Best Local Similarity 26.9%; Pred. No. 1.3e-16;
Matches 97; Conservative 63; Mismatches 135; Indels 66; Gaps 13;

QY      30 YVEMTVGSPQTLNLTIVDTGSSNFAVGAAPHPF-----LRRYQRLSSTYRDLRK 82
Db      75 YFGTISGTTPPDFTYVFTGSSNLMV--PSTYCSSLACALHKKRNPEDSSTYQGTSTF 131

```



```
OY 119 TESDK-FEINGSNMEGILGLAYAEIARPPDDSLBEPFDSLVRKQTHV-PNLFSLQLCGAGFP 176
Db 166 SETEPGSEFLYARAPDGLGLAYPSTIS--SSGATPVFDNIMWQGLVSODLFSVYLAD--- 220
OY 177 LNQSEVLASVGGSMITIGIDHSLYTGSLMTPIREMYEYIIVREINQODLKMCKEY 236
Db 221 -DQS-----GSVVIFGIDSSYTYTSLMWVPVSVEGWQISVDSTIMNGEALA--CAE- 270
OY 237 NYDKSIYDSGTFNTLRPKKVEAAVKSIIKAASSTEFKPPDGFWLGEOLV-CWQAGTTPWNI 295
Db 271 -GCCQIVDTGTSLSLTGPTSPIANIOSDIGASEND-----GEMVWSCSAISSLPDIY 321
OY 296 FPVLSLYLMGEVTNQSFRITLPOQYLKRPVEDVATSQDDCYK-----FAISQSTGYMG 350
Db 322 F-----TINGIQYPVPPSAY-----ILOSQSGSCISFGQMDVPTESGSELMWILG 364
OY 351 AVIMEGFYVFPDRAKKRIGFA 371
Db 365 DVEIRQYFTVEDRANNOVOLA 385
```

Search completed: August 7, 2002, 09:14:17
Job time: 41 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2002, 09:13:35 ; Search time 22.58 seconds

(without alignments)
493.272 Million cell updates/sec

Title: US-09-724-571-43

Perfect score: 2419

Sequence: 1 ETDEEPEEPGRSGFVEMVD.....CLNCLRQOHDFADISLKLK 456

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA.*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCrus.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2413	99.8	501	4	US-09-009-191-2
2	2320	95.9	774	4	US-09-009-191-4
3	1171.5	48.4	518	3	US-08-999-723-2
4	1171.5	48.4	518	4	US-09-434-427-2
5	1146.5	47.4	514	4	US-09-717-432-2
6	1146.5	47.4	514	4	US-09-912-484-2
7	299.5	12.4	396	1	US-08-208-007A-13
8	299.5	12.4	396	4	US-09-032-523-9
9	295.5	12.2	412	1	US-08-208-007A-12
10	295.5	12.2	412	4	US-08-974-691-4
11	279.5	11.6	456	6	5217891-15
12	273.5	11.3	409	1	US-09-640-305-6
13	273.5	11.3	409	1	US-08-360-673-6
14	273.5	11.3	427	2	US-08-846-021A-8
15	270	11.2	410	1	US-08-088-633-2
16	270	11.2	410	1	US-08-245-756-2
17	270	11.2	410	1	US-08-441-750-2
18	270	11.2	410	2	US-08-441-751-2
19	270	11.2	410	5	PCT-US92-02521-2
20	250	10.3	349	4	US-09-032-523-3
21	227	9.4	398	1	US-08-328-314-2
22	227	9.4	398	4	US-08-731-045-2
23	212	8.8	419	4	US-08-974-691-3
24	210	8.7	397	3	US-09-079-415-2
25	196.5	8.1	430	1	US-08-535-237-2
26	195.5	8.1	427	1	US-07-958-222A-2
27	193	8.0	330	3	US-08-115-753-1

28	193	8.0	419	3	US-08-115-753-2	Sequence 2, Appl1
29	193	8.0	419	3	US-08-115-753-3	Sequence 33, Appl1
30	187	7.7	420	4	US-09-008-211A-4	Sequence 4, Appl1
31	187	7.7	420	4	US-08-974-691-8	Sequence 8, Appl1
32	184.5	7.6	395	1	US-08-723-938-3	Sequence 3, Appl1
33	184.5	7.6	395	2	US-09-080-538-3	Sequence 3, Appl1
34	184	7.6	445	4	US-08-974-691-6	Sequence 6, Appl1
35	184	7.6	451	4	US-08-974-691-2	Sequence 2, Appl1
36	149	6.2	437	4	US-09-353-332-2	Sequence 2, Appl1
37	128.5	5.4	140	3	US-09-211-631-13	Sequence 13, Appl1
38	128.5	5.4	140	4	US-09-265-628-13	Sequence 13, Appl1
39	129.5	5.4	140	4	US-09-001-141-11	Sequence 11, Appl1
40	129.5	5.4	140	4	US-09-532-803-6	Sequence 6, Appl1
41	129.5	5.4	140	4	US-09-653-403-14	Sequence 14, Appl1
42	97	4.0	1030	4	US-09-091-117-2	Sequence 2, Appl1
43	95.5	3.9	280	4	US-09-160-246-14	Sequence 14, Appl1
44	92.5	3.8	1097	2	US-08-680-326-39	Sequence 39, Appl1
45	88	3.6	746	2	US-08-838-219B-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
Sequence 2, Application US/09009191
Patent No. 6319689
GENERAL INFORMATION:
APPLICANT: POWELL, DAVID
APPLICANT: CHAPMAN, CONRAD
APPLICANT: MURPHY, KAY
APPLICANT: SMITH, TRUDI
TITLE OF INVENTION: ASP2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,191
FILING DATE: 20-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 9701684.4
FILING DATE: 28-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70368
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-191-2
Query Match 99.8%; Score 2413; DB 4; Length 501;
Best Local Similarity 99.8%; Pred. No. 3.3e-251;

Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEPEEPGRGSRFVEMVNDLGRKSGGGYVEMTVGSPQTLNLLVDTGSSNEFAVGAP 60
 DB 46 EDEDEPEEPGRGSRFVEMVNDLGRKSGGGYVEMTVGSPQTLNLLVDTGSSNEFAVGAP 105
 QY 61 HPELHRYTOROLSTYRDLRKGVYPTYGKWEGLGTDLVSIIPHGNVTVRANIAATTE 120
 DB 106 HPELHRYTOROLSTYRDLRKGVYPTYGKWEGLGTDLVSIIPHGNVTVRANIAATTE 165
 QY 121 SSKFPTNSNMGGIIGLVAETARPDSDLPEPFDLVKQTHVNPFLSLOLCGAGFPLNOS 180
 DB 166 SSKFPTNSNMGGIIGLVAETARPDSDLPEPFDLVKQTHVNPFLSLOLCGAGFPLNOS 225
 QY 181 EYLVASVSGSMIIGIDHSIYTGSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDK 240
 DB 226 EYLVASVSGSMIIGIDHSIYTGSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDK 285
 QY 241 SYVDGTTNLRPKVFEAAVKSIAASTEKFPDGFMLGEOLVCMWAGTTPMNIPEVIS 300
 DB 286 SYVDGTTNLRPKVFEAAVKSIAASTEKFPDGFMLGEOLVCMWAGTTPMNIPEVIS 345
 QY 301 LYLMGEVNTQSFRTITLPOQYLRPVEDVATSDDCYKFAISQSGTGVWGAIVMEGFYV 360
 DB 346 LYLMGEVNTQSFRTITLPOQYLRPVEDVATSDDCYKFAISQSGTGVWGAIVMEGFYV 405
 QY 361 FPRARRIGFAVSACHVDEFTAAVEGFVTLDMEDCGYNIPQDESTLMTIAYVMAI 420
 DB 406 FPRARRIGFAVSACHVDEFTAAVEGFVTLDMEDCGYNIPQDESTLMTIAYVMAI 465
 QY 421 CALFMLPLCLMVCORCLRCRQOHDDFADDISLKL 456
 DB 466 CALFMLPLCLMVCORCLRCRQOHDDFADDISLKL 501

RESULT 2
 US-09-009-191-4
 ; Sequence 4, Application us/09009191
 ; Patent No. 6319689
 ; GENERAL INFORMATION:
 ; APPLICANT: POWELL, DAVID
 ; APPLICANT: CHAPMAN, CONRAD
 ; APPLICANT: MORPHY, KAY
 ; APPLICANT: SMITH, TRUDI
 ; TITLE OF INVENTION: ASP2
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: RATNER & PRESTIA
 ; STREET: P. O. BOX 980
 ; CITY: VALLEY FORGE
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19482
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/009,191
 ; FILING DATE: 20-JAN-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: UK 9701684.4
 ; FILING DATE: 28-JAN-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PRESTIA, PAUL F
 ; REGISTRATION NUMBER: 23,031
 ; REFERENCE/DOCKET NUMBER: GH-70368
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-407-0700
 ; TELEFAX: 610-407-0701
 ; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 774 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-009-191-4

Query Match 95.9%; Score 2320; DB 4; Length 774;
 Best Local Similarity 98.9%; Pred. No. 7,1e-241;
 Matches 439; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 GSFEVMDVNLGKSGGGYVEMTVGSPQTLNLLVDTGSSNEFAVGAPHPPLHRYTOROL 72
 DB 1 GSFEVMDVNLGKSGGGYVEMTVGSPQTLNLLVDTGSSNEFAVGAPHPPLHRYTOROL 60
 QY 73 SSTYDRLKRGYVPTYGKWEGLGTDLVSIIPHGNVTVRANIAATTESDKFPTNSNME 132
 DB 61 SSTYDRLKRGYVPTYGKWEGLGTDLVSIIPHGNVTVRANIAATTESDKFPTNSNME 120
 QY 133 GILGLAVALIARPDSDLPEPFDLVKQTHVNPFLSLOLCGAGFPLNOS EYLVASVGSMTI 192
 DB 121 GILGLAVALIARPDSDLPEPFDLVKQTHVNPFLSLOLCGAGFPLNOS EYLVASVGSMTI 180
 QY 193 GGIDHSIYTGSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDKSIYDSGTTNLR 252
 DB 181 GGIDHSIYTGSLWYTPIRREMYEVIIVVEINGODLKMCKEYNDKSIYDSGTTNLR 240
 QY 253 PKKVEEAAVKSIAASTEKFPDGFMLGEOLVCMWAGTTPMNIPEVISLYLMGEVNTQSF 312
 DB 241 PKKVEEAAVKSIAASTEKFPDGFMLGEOLVCMWAGTTPMNIPEVISLYLMGEVNTQSF 300
 QY 313 RTITLPOQYLRPVEDVATSDDCYKFAISQSGTGVWGAIVMEGFYVDEARRKIGFV 372
 DB 301 RTITLPOQYLRPVEDVATSDDCYKFAISQSGTGVWGAIVMEGFYVDEARRKIGFV 360
 QY 373 SACHVDEFTAAVEGFVTLDMEDCGYNIPQDESTLMTIAYVMAICALFMLPLCLMV 432
 DB 361 SACHVDEFTAAVEGFVTLDMEDCGYNIPQDESTLMTIAYVMAICALFMLPLCLMV 420
 QY 433 COMRCLRCRQOHDDFADDISLKL 456
 DB 421 COMRCLRCRQOHDDFADDISLKL 444

RESULT 3
 US-08-999-723-2
 ; Sequence 2, Application US/08999723A
 ; Patent No. 6025180
 ; GENERAL INFORMATION:
 ; APPLICANT: Powell, David J.
 ; APPLICANT: Southan, Christopher
 ; APPLICANT: Chapman, Conrad G.
 ; APPLICANT: Evans, Joanne R.
 ; TITLE OF INVENTION: ASP1
 ; FILE REFERENCE: GH70262
 ; CURRENT APPLICATION NUMBER: US/08/999,723A
 ; CURRENT FILING DATE: 1997-10-06
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 518
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-08-999-723-2

Query Match 48.4%; Score 1171.5; DB 3; Length 518;
 Best Local Similarity 49.6%; Pred. No. 2,2e-117;
 Matches 222; Conservative 80; Mismatches 139; Indels 7; Gaps 4;

Db 362 NASSSFRTLLPOLYLQPMGAGFN - ECRFGISSNTNLVIGATVEGEFYVFPDQR 420
QY RIGRNASCHHDEFRRAVGPVTLDMDCGGINIQDTESTLMTAYMAALICAFML 426
QY RIGRNASCHHDEFRRAVGPVTLDMDCGGINIQDTESTLMTAYMAALICAFML 426
Db 421 RGRNVSFCALIEGTYSSEISGPFSTEDIASNCYPAQALNEPLIMVSYALMSVCGATLL 480
QY 427 PLCLTAV 432
Db 481 VLITLL 486

```

, RESULT 6
, US-09-912-484-2
, Sequence 2, Application US/09912484
, Patent No. 6358725
, GENERAL INFORMATION:
, APPLICANT: Christie, Gary
, APPLICANT: Li, Xiaotong
, APPLICANT: Powell, David J.
, APPLICANT: Zhu, Yuan
, TITLE OF INVENTION: Mouse Aspartic Secretase-1 (MASP1)
, FILE REFERENCE: GP-70663-D1
, CURRENT APPLICATION NUMBER: US/09/912,484
, CURRENT FILING DATE: 2001-07-25
, PRIOR APPLICATION NUMBER: 60/166,974
, PRIOR FILING DATE: 1999-11-23
, PRIOR APPLICATION NUMBER: 09/717,432
, PRIOR FILING DATE: 2000-11-21
, NUMBER OF SEQ ID NOS: 2
, SOFTWARE: FastSeq for Windows Version 3.0
, SEQ ID NO 2
, LENGTH: 514
, TYPE: PRF
, ORGANISM: MUS MUSCULUS
, US-09-912-484-2

```

Query Match	47.4%;	Score 1146.5;	DB 4;	Length 514;
Best Local Similarity	50.7%;	Pred. No. 1.1e-114;		
Matches 216;	Conservative 75;	Mismatches 130;	Indels 5;	Gaps 3

[illegible]

```

1      RESULT 7
2      US-08-208-007A-13
3      ; Sequence 13, Application US/08208007A
4      ; Patent No. 5501969
5      GENERAL INFORMATION:
6      APPLICANT: HASTINGS, ET AL.
7      TITLE OF INVENTION: Human Osteoclast-Derived Cathepsin
8      NUMBER OF SEQUENCES: 14
9      CORRESPONDENCE ADDRESSES:
10     ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
11     ADDRESSEE: CECCHI, STEWART & OLSTEIN
12     STREET: 6 BECKER FARM ROAD
13     CITY: ROSELAND
14     STATE: NEW JERSEY
15     COUNTRY: USA
16     ZIP: 07068
17     COMPUTER READABLE FORM:
18     MEDIUM TYPE: 3.5 INCH DISKETTE
19     COMPUTER: IBM PS/2
20     OPERATING SYSTEM: MS-DOS
21     SOFTWARE: WORD PERFECT 5.1
22     CURRENT APPLICATION DATA:
23     APPLICATION NUMBER: US/08/208,007A
24     FILING DATE: March 8, 1994
25     CLASSIFICATION: 435
26     PRIOR APPLICATION DATA:
27     APPLICATION NUMBER: No. 5501969e
28     FILING DATE: No. 5501969e
29     ATTORNEY/AGENT INFORMATION:
30     NAME: FERRARO, GREGORY D.
31     REGISTRATION NUMBER: 36,134
32     REFERENCE/DOCKET NUMBER: 325800-95
33     TELECOMMUNICATION INFORMATION:
34     TELEPHONE: 201-994-1700
35     TELEFAX: 201-994-1744
36     INFORMATION FOR SEQ ID NO: 13:
37     SEQUENCE CHARACTERISTICS:
38     LENGTH: 396 AMINO ACIDS
39     TYPE: AMINO ACID
40     STRANDNESS: ACID
41     TOPOLOGY: LINEAR
42     MOLECULE TYPE: PROTEIN
43     US-08-208-007A-13

```

Query Match	12.4%;	Score 299.5;	DB 1;	Length 396;
Best Local Similarity	25.9%;	Pred. No. 1.3e-23;		
Matches 100;	Conservative	65;	Mismatches 148;	Indels 73;
				Gaps 16

[illegible]

QY 347 -TVMGAVIMEGFYVVEDRARRKIGFA 371
DB 367 LMIIGDVFIROFYSVFDKGNRRVGLA 392

RESULT 8
US-09-032-523-9

; Sequence 9, Application US/09032523
; Patent No. 6232454
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Cortley, Neil C.
; APPLICANT: Guegler, Karl
; APPLICANT: Baugh, Mariah
; TITLE OF INVENTION: HUMAN PROTEINASE MOLECULES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,523
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0479 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 396 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 181994
; US-09-032-523-9

Query Match 12.4%; Score 299.5; DB 4; Length 396;
Best Local Similarity 25.9%; Pred. No. 1.3e-23;
Matches 100; Conservative 65; Mismatches 148; Indels 73; Gaps 16;

QY 3 DEPEEPGRGSFVEMVNLNRKSGQGYVEMTVGSPQTLNIVDTGSSNFAVGA----- 58
DB 63 DOSAKEP-----LINVLD-----MEYFGTISTISPPQNTVIFDTGSSNLWVPSYCT 110
QY 59 AHPRLHYRQRLSTYRDLRKGYVYPTQGWEGELGTLVSIPIGPNVYRANIAAI 118
DB 111 SPACTHRRFPQSSSTYSQPGQSFSDYGTGSGISGIIAGDQVSV-BGLTVVGGQGFESV 169
QY 119 TESDFKFNWMEGIIILAVAEIARPDLSLEPPFDSLVKOTHPNLFSLQDLCAGFPLN 178
DB 170 TEPGQTFVD-AEFDGIIIGLGPISLA--VGVVTPVFDNNMAO-----NLVDLPMSFSVYSSN 222

QY 179 QSEVLASVGSMIIGIDHSITGSLNTPPIREMYEYIIVREINODLMDCKEYNY 238
DB 223 PE---GGASSELPFGYDHSFSGSLNWPVTKQAYWQJALDNIQVGG--TVMFCSE--G 275
QY 239 DRSIVDSGTNLRPKKVEAAVKSIAKASSTKPPDGFMLGEOQVCMQAGTTPNINIPV 298
DB 276 COAIVDTGTSILTGPDSKIKQLOINAIQAAP-----VDGEYAVE-----CANLNVMPD 322
QY 299 ISILYLMGEVTNOSFRITLLPQOYLRPVEDVATSDQCYKFAISQSTG----- 346
DB 323 VFTTNG-----VPYTLSPYAV--TLDFVDGMQFC-----SSGFGDGIHPAPG 366
QY 347 -TVMGAVIMEGFYVVEDRARRKIGFA 371
DB 367 LMIIGDVFIROFYSVFDKGNRRVGLA 392

RESULT 9
US-08-208-007A-12

; Sequence 12, Application US/08208007A
; Patent No. 5501969
; GENERAL INFORMATION:
; APPLICANT: HASTINGS, ET AL.
; TITLE OF INVENTION: Human Osteoclast-Derived Cathepsin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESS: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/208,007A
; FILING DATE: Match 8, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: No. 5501969e
; FILING DATE: No. 5501969e
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-208-007A-12

Query Match 12.2%; Score 295.5; DB 1; Length 412;
Best Local Similarity 28.5%; Pred. No. 3.8e-23;
Matches 105; Conservative 58; Mismatches 146; Indels 59; Gaps 15;

QY 30 YVEMTVGSPQTLNIVDTGSSNFAVGAAPHEL-----HRYRQRLSTYRDLRKGV 83
DB 79 YVEIGIGTPPOCFVTVPDTGSSNLWVPSIHCXLDIACMIHKKVYKSDKSTYYVKNQTSF 138
QY 84 YVPYTOGWEGELGTLVSIPI-----HGPNTYRANIAAIRESDFKFNWMEGI 134
DB 139 DIHYGSGSLSGTISQDTVAVPCOSASASALGVKVERQVFGEATRKQPGITFIKAFDGI 198

QY 135 LGLAVALIARPDLSLEPFDSLVKQTHV-PNLSLQCGAGFPLNQSSEVLASVGSMTIG 193
 Db 199 LGMAVPRIS--VNNVLPVFNLMQKLVQNIISFYI-----SRDPAQPGGELMIG 248
 QY 194 GIDHSLYTGLWYPIRREWEYEVIIIVREI-NGQDLKMDCKEYNDKSIYDSGTNRL 252
 Db 249 GTDSKYKGSLSYLVNTRKAVQVHLDQVEVASGLTL--CKE--GCEAIVDTGTSIMVG 303
 QY 253 PKKVEAAVKSIRKASSTKPPDFWMLGEOLV-CMOAGTTPMNIFFVISLYLMGEVTNOS 311
 Db 304 PVDEVRELQKALGAVPLIQ-----GEYMIPEKYST-----LPAITLKLGG---KG 346
 QY 312 FRITILPOOYLRPVEDVATSODDCYKFAISQ-----SSTGTWGANVMEGFYVVEDRARK 366
 Db 347 YKLS--PEDYTLKVSQAKTL--CLSGFMGMDIPPSGPFMLIGDVFIGRYTVFDRDNN 402
 QY 367 RIGFAVSA 374
 Db 403 RVGFAEAA 410

RESULT 10
 US-08-974-691-4
 ; Sequence 4, Application US/08974691
 ; Patent No. 6225103

GENERAL INFORMATION:
 APPLICANT: Keolsch, Gerald

APPLICANT: Lin, Xinli

APPLICANT: Tang, Jordan

TITLE OF INVENTION: Cloning and Characterization of Napsin

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst

STREET: 2800 One Atlantic Center, 1201 W. Peachtree

CITY: Atlanta

STATE: GA

COUNTRY: USA

ZIP: 30309-3450

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,691

FILING DATE: 20-NOV-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/031,196

FILING DATE: 20-NOV-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/046,126

FILING DATE: 09-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: OMRF 166

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-873-8794

TELEFAX: 404-873-8795

INFORMATION FOR SEQ. ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 412 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

US-08-974-691-4

Query Match 12.2%; Score 295.5; DB 4; Length 412;
 Best Local Similarity 28.5%; Pred. No. 3.8e-23;
 Matches 105; Conservative 58; Mismatches 146; Indels 59; Gaps 15;

QY 30 YVEVMTVSGPPQTLNIVDTGSSNFAVGAAPPEL-----HRYQROLSTYRLRGV 83
 Db 79 YVEIIGIGIPPOCFVFPDGSNLMWPSIHCKLDICMIHKYNSDKSSYVNGTSF 138
 QY 84 YVPTQKKEGELGDTLVSP-----HGPNTVIRANIAITSDKFFINGSNMGEI 134
 Db 139 DIHYGGSLSGYLSQDTVSVCOSASSALGVKVERVFEARIKOPIITIAKFDGI 198
 QY 135 LGLAVALIARPDLSLEPFDSLVKQTHV-PNLSLQCGAGFPLNQSSEVLASVGSMTIG 193
 Db 199 LGMAVPRIS--VNNVLPVFNLMQKLVQNIISFYI-----SRDPAQPGGELMIG 248
 QY 194 GIDHSLYTGLWYPIRREWEYEVIIIVREI-NGQDLKMDCKEYNDKSIYDSGTNRL 252
 Db 249 GTDSKYKGSLSYLVNTRKAVQVHLDQVEVASGLTL--CKE--GCEAIVDTGTSIMVG 303
 QY 253 PKKVEAAVKSIRKASSTKPPDFWMLGEOLV-CMOAGTTPMNIFFVISLYLMGEVTNOS 311
 Db 304 PVDEVRELQKALGAVPLIQ-----GEYMIPEKYST-----LPAITLKLGG---KG 346
 QY 312 FRITILPOOYLRPVEDVATSODDCYKFAISQ-----SSTGTWGANVMEGFYVVEDRARK 366
 Db 347 YKLS--PEDYTLKVSQAKTL--CLSGFMGMDIPPSGPFMLIGDVFIGRYTVFDRDNN 402
 QY 367 RIGFAVSA 374
 Db 403 RVGFAEAA 410

RESULT 11
 5217891-15
 ; Patent No. 5217891
 ; APPLICANT: BRAKE, ANTHONY J.; VAN DEN BERG, JOHAN A.
 ; TITLE OF INVENTION: DNA CONSTRUCTS CONTAINING A KLUVEROMYCES
 ; A FACTOR LEADER SEQUENCE FOR DIRECTING SECRETION OF HETEROLOGOUS
 ; POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 23
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/507,398
 ; FILING DATE: 09-APR-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 78,551
 ; FILING DATE: 28-JUL-1987
 ; SEQ. ID NO. 15:
 ; LENGTH: 458
 ; 5217891-15

Query Match 11.6%; Score 279.5; DB 6; Length 458;
 Best Local Similarity 26.8%; Pred. No. 2.4e-21;
 Matches 95; Conservative 67; Mismatches 128; Indels 65; Gaps 15;

QY 31 YVEVMTVSGPPQTLNIVDTGSSNFAVGAAPPEL-----HRYQROLSTYRLRGV 83
 Db 152 YFGIYGTGPEPFVLEFDGSSDFWV--PSITYCKSNACKNHQRPDOKSSIFOMLGDL 208
 QY 84 YVPTQKKEGELGDTLVSP-----HGPNTVIRANIAITSDKFFINGSNMGEI 134
 Db 209 SIHYGTGSMOGLIGYTVVSNVLDIQVGLSTGPGVF--TYAEFDIGMAVPSIA 266
 QY 144 RPDLSLEPFDSLVKQTHV-PNLSLQCGAGFPLNQSSEVLASVGSMTIGIDHSITG 202
 Db 267 --SEYSLPVVDNMMNRHLVAQDLFSYMDRNG--QESMLT-----LGAIDPSYTG 313
 QY 203 SLWYTPPIRREWEYEVIIIVREIINGQDLKMD--CKEYNDKSIYDSGTNRLPKKVEAA 260
 Db 314 SLHWVPVTVQOYWOFTVDSVLTISGVVACEGCG-----QALIDTGTSLVGPSSDILNI 367

QY 261 VKSIKASTKFEKFDGFWLGE-QLVCMQAGTTPMNIFFVISLYLMGEVYNOSFRITLIPQ 319
DB 368 00A1GATGNO-----YGFEDIDCDNLSTYMTVVF-----ELNCKMYLT--PS 408
QY 320 QYLREVEDVATSDDCY---KFAISQSTGTVMGAVIMEGTVVDFDRARKRIGFA 371
DB 409 AY-----TSQDGFCTSGFQSEHNSQGMILGDVFEIRYYSVDFRANMLVGLIA 455

RESULT 12
US-09-640-305-6
Sequence 6, Application US/09640305
Patent No. RE37447
GENERAL INFORMATION:
APPLICANT: Fleet, Reinhard
Fournier, Alain
Yeh, Patrice
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
PREPARATION AND USE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/640,305
APPLICATION NUMBER: US/09/640,305
FILING DATE: 16-Aug-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/360,673
FILING DATE: 06-FEB-1995
APPLICATION NUMBER: WO PCT/FR93/00623
FILING DATE: 23-JUN-1993
APPLICATION NUMBER: FR 92/07785
FILING DATE: 25-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST92040-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-640-305-6

Query Match 11.3%, Score 273.5; DB 1; Length 409;
Best Local Similarity 25.3%; Pred. No. 8.7e-21;
Matches 92; Conservative 59; Mismatches 144; Indels 69; Gaps 11;

QY 30 YYEVTGSPPTQTLNLTDTGSSNFAVGAAP---HPLHRYRQRLSSTYRDLRKGYV 85
DB 96 YTEITLTSPOSSEKVIIDTGSNLMWVSABCGSIACFLHRTKDYHSASTYKANGSEFAI 155
QY 86 PYTGKMBEGLDGLVSIPIHGPNTVTRANIAITESDKFFITGNSMWEIGLGLAETARP 145
DB 156 QYSGSLGEGYSDLLTI--GDVLPDODFAEATSEBGLAFAPKFPDGLGLADISIS-V 212
QY 146 DDLSEPFDSLKQTHVNLFSLQLCGAGFLNGLSEVLASVGGSMITGGIDHSIYTGSLW 205

DB 213 NRIYPPVYNAIKKLDDPVA-----FYLGDSD-KSEDEGEASFGGIDEKTYGETI 264
QY 206 YPIREMYEYVLIYVETNGDLMKDCKEYNYKDSIYSGTNTLRPKKVEAIVKSIK 265
DB 265 WLPVRRKAYWE-----VKEGIGLGEYATLTGHRGNAIDTGISIALPSGLAEIILAEIG 319
QY 266 AASSTKFPDGGWLGELQVCMQAGTTPMNIFFVISLYLMGEVYNOSFRIT----- 315
DB 320 AKKG-----WSQYSVDCESRDS---LPDLTLNFNG---YNFTITAVDTLEVSG 363
QY 316 -----ILPQYLREVEDVATSDDCYKFAISQSTGTVMGAVIMEGTVVDFDRARKRIGF 370
DB 364 SCISAFTPMDFPEYGPPLA-----ITGAFILKRYSIYDIGHDAVGL 405
QY 371 AVSA 374
DB 406 AKAA 409

RESULT 13
US-08-360-673-6
Sequence 6, Application US/08360673
Patent No. 5679544
GENERAL INFORMATION:
APPLICANT: Fleet, Reinhard
Fournier, Alain
Yeh, Patrice
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
PREPARATION AND USE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002

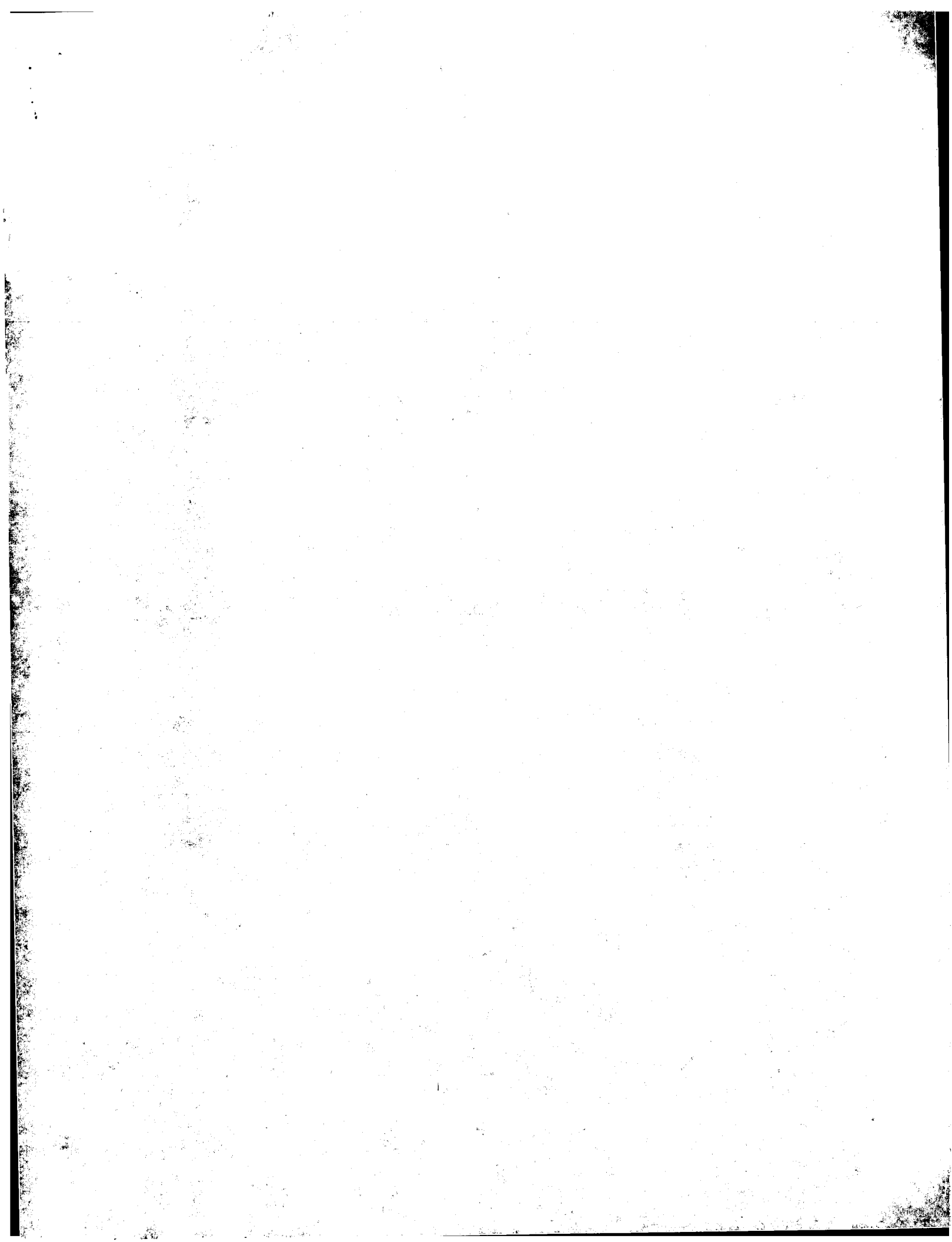
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,673
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR93/00623
FILING DATE: 23-JUN-1993
APPLICATION NUMBER: FR 92/07785
FILING DATE: 25-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: ST92040-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-360-673-6

Query Match 11.3%, Score 273.5; DB 1; Length 409;
Best Local Similarity 25.3%; Pred. No. 8.7e-21;
Matches 92; Conservative 59; Mismatches 144; Indels 69; Gaps 11;

QY 30 YYEVTGSPPTQTLNLTDTGSSNFAVGAAP---HPLHRYRQRLSSTYRDLRKGYV 85


```
Search completed: August 7, 2002, 09:14:48
Job time: 73 sec
```



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2002, 09:13:35 ; Search time 56.55 Seconds

(without alignments)
895.661 Million cell updates/sec

Title: US-09-724-571-43

Perfect score: 2419
Sequence: 1 ETDEEPEEPEGRGSEFVEMVD.....CLRLRQOHDFADISILK 456Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

_A.Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2419	100.0	456	21	AA07897
2	2419	100.0	488	22	AAB6572
3	2419	100.0	488	22	AAB6134
4	2419	100.0	501	21	AA94767
5	2419	100.0	501	21	AA07896
6	2419	100.0	503	22	AAB6573
7	2419	100.0	503	22	AAB6135
8	2414	99.8	501	21	AA58425
9	2414	99.8	501	22	AAE10629
10	2414	99.8	501	22	AAE06859
11	2414	99.8	501	22	AAU06603

12	2414	99.8	501	22	AAU07202	Human aspartyl pro
13	2414	99.8	501	22	AAE02581	Human aspartyl pro
14	2413	99.8	501	19	AAW59807	Amino acid sequenc
15	2392	98.9	501	21	AA94769	Rat beta-secretase
16	2390	98.8	501	21	AA94768	Murine beta-secret
17	2390	98.8	501	21	AAV88427	Murine aspartyl pr
18	2390	98.8	501	22	AAE10631	Murine aspartyl pr
19	2390	98.8	501	22	AAE06861	Murine aspartyl pr
20	2390	98.8	501	22	AAU06605	Mouse Aspartyl pro
21	2390	98.8	501	22	AAU07204	Mouse aspartyl pro
22	2390	98.8	501	22	AA84948	Mouse aspartyl pro
23	2390	98.8	501	22	AAE02583	Mouse aspartyl sec
24	2351.5	97.2	969	22	ABG09611	Novel human diago
25	2320	95.9	790	19	AAW59808	Partial amino acid
26	2264.5	93.6	476	21	AAV88426	Human aspartyl pro
27	2264.5	93.6	476	22	AAE10630	Human aspartyl pro
28	2264.5	93.6	476	22	AAE06860	Human aspartyl pro
29	2264.5	93.6	476	22	AAU06604	Human aspartyl pro
30	2264.5	93.6	476	22	AAU07203	Human aspartyl pro
31	2264.5	93.6	476	22	AAE02582	Human aspartyl pr
32	2238.5	92.5	476	22	AAE06909	Murine aspartyl pr
33	2238.5	92.5	476	22	AAE06919	Murine aspartyl pr
34	2212	91.4	427	22	AA93866	Human polypeptide,
35	2160	89.3	453	21	AAV88438	Modified human asp
36	2160	89.3	453	22	AAE10642	Human-Asp 2(a) del
37	2160	89.3	453	22	AAE06872	Human-Asp 2(a) pro
38	2160	89.3	453	22	AAU06616	Human-Asp 2(a) del
39	2160	89.3	453	22	AAU07215	Human aspartyl pro
40	2160	89.3	453	22	AAE02594	Human-Asp 2(a) del
41	2160	89.3	459	21	AAV88439	Modified human asp
42	2160	89.3	459	22	AAE10643	Human-Asp 2(a) pro
43	2160	89.3	459	22	AAE06873	Human-Asp2(a) del
44	2160	89.3	459	22	AAU06617	Human-pro-Asp 2(a)
45	2160	89.3	459	22	AAU07216	Human aspartyl pro

ALIGNMENTS

RESULT 1

ID AAB07897 standard; Protein; 456 AA.

AC AAB07897;

DT 14-NOV-2000 (first entry)

DE Active enzyme portion of human beta-secretase enzyme.

KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;

KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;

KW inhibitor; ss.

OS Homo sapiens.

PN WO20047618-A2.

PD 17-AUG-2000.

PF 10-FEB-2000; 2000WO-US03819.

PR 10-FEB-1999; 99US-0119571.

PR 15-JUN-1999; 99US-0139172.

PA (ELAN) ELAN PHARM INC.

PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;

PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;

DR WPI: 2000-533011/48.

PT Purified beta-secretase protein used in assays to discover inhibitors

PT which can be used for the treatment of amyloidogenic diseases e.g.

PT Alzheimer's disease -
 XX
 PS Claim 24; Fig 2b; 121pp; English.
 XX
 CC The specification describes a beta-secretase enzyme. The enzyme cleaves
 CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
 CC enzyme is therefore implicated in the production of amyloid plaque
 CC components which accumulate in the brains of individuals afflicted with
 CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
 CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
 CC disease-like pathology to test if they maintain or improve cognitive
 CC ability or reduce the plaque burden. The compounds are used for the
 CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
 CC present sequence represents the active enzyme portion of human
 CC beta-secretase enzyme.
 CC
 XX
 SQ Sequence 456 AA;

Query Match 100.0%; Score 2419; DB 21; Length 456;
 Best Local Similarity 100.0%; Pred. No. 3.9e-242; Indels 0; Gaps 0;
 Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEPEEPGRGRGSFVEMVDNLKSGSGGYVEMTVGSPPTNLILVDTGSSNFAVGAP 60
 DB 1 etdeepeepgrgrgsfvevmdnlrgsksggyvemtvgspptnlilvdtgssnfavgap 60
 QY 61 HPFLHRYQROLSTYRDLRKGVYVPTQKWEGLGTDLVSIHPGPNTVANAIAITE 120
 DB 61 hpflhryqrolstyrdlrkgyvypytqkwegelgtldvlsihpgpntvanaiaalte 120
 QY 121 SDKFFINGSNMGGIIGLVAEIAEPDSDLPEPFDLSVKOTHPNFIQLCGAGPELNS 180
 DB 121 sdkffingsnmggiiglvaeiaepdsslepfdsivkothpvnfislcgagpflngs 180
 QY 181 EYLAASVGSMTIIGIDHSLYTGSLWYTPIRREMYEVIIIVRVEINGODIKMCKEYNDK 240
 DB 181 eylvasvgsmtiigidhsltytgslwytptirremyeviiivrveingodikmckeynydk 240
 QY 241 SYVDGSTNLRLPKKVEFAVKSIIKAASSTKEKPPDGFWMGEOLVCOMOAGTTPMNIPEVVS 300
 DB 241 syvdgstnlrlpkkvefaavksikaasstekpdpdgfwlgeqlvcwgaqctpwnlfpvvs 300
 QY 301 LYIMGEVNTQSFRTITLPQOYLRPVEDVATSDDCYKFAISOSTGTWGAIVMEGFYVV 360
 DB 301 lyimgevtngsfrtitlppqylrpvedvatsddcykfaissgstgtwmgavimegfyvv 360
 QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPPVTLDMEDCGYNIPQDESTLMTIAYMAAI 420
 DB 361 fdrarkrigfavsachvhdetrtaavegpfvtldmedcgynipqdestlmtiayymaa 420
 QY 421 CALFMLPICLMWCQWRCLRCLRQOHDDFADDISILK 456
 DB 421 calfmplclmwcwrcrlrclrqghddfaddisilk 456

RESULT 2
 AAB66572
 ID AAB66572 standard; Protein; 488 AA.

AC AAB66572;

DF 12-APR-2001 (first entry)

DE Human memapsin 2.

XX Human; memapsin 2; nootropic; neuroprotective; amyloid precursor protein;

KW APP; memapsin 2 inhibitor; Alzheimer's disease.

OS Homo sapiens.

XX WO200100665-A2.

XX

PD 04-JAN-2001.
 XX
 XX 27-JUN-2000; 2000WO-US17742.
 PE
 XX 28-JUN-1999; 9905-0141363.
 PR 30-NOV-1999; 9905-0168060.
 PR 25-JAN-2000; 2000US-0177836.
 PR 27-JAN-2000; 2000US-0178368.
 PR 08-JUN-2000; 2000US-0210292.
 XX
 PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
 PA (UNIT) UNITV ILLINOIS FOUND.
 PI Tang JUN, Hong L, Ghosh AK;
 XX
 DR WPI: 2001-137933/14.
 DR N-PSDB; AAF31848.

PT Novel memapsin 2 inhibitors which bind to active site of memapsin 2
 PT having 2 catalytic aspartic residues and substrate binding cleft, used
 PT to treat Alzheimer's disease by blocking amyloid precursor protein
 PT cleavage

PS Example 1; Page 72-74; 86pp; English.

CC The present sequence is given in a specification relating to an inhibitor
 CC of catalytically active memapsin 2. The inhibitor binds to the memapsin 2
 CC active site, which is defined by the presence of two catalytic aspartic
 CC residues and a substrate binding cleft. The inhibitor is useful for
 CC the treatment and diagnosis of Alzheimer's disease. It is useful in
 CC screens for individuals with a genetic predisposition to Alzheimer's
 CC disease. The inhibitor is useful as a reagent for specifically binding to
 CC memapsin 2 or memapsin 2 analogues and for aiding in memapsin 2
 CC isolation, purification and characterisation.

XX Sequence 488 AA;

Query Match 100.0%; Score 2419; DB 22; Length 488;
 Best Local Similarity 100.0%; Pred. No. 4.3e-242;
 Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEPEEPGRGRGSFVEMVDNLKSGSGGYVEMTVGSPPTNLILVDTGSSNFAVGAP 60
 DB 33 etdeepeepgrgrgsfvevmdnlrgsksggyvemtvgspptnlilvdtgssnfavgap 92
 QY 61 HPFLHRYQROLSTYRDLRKGVYVPTQKWEGLGTDLVSIHPGPNTVANAIAITE 120
 DB 93 hpflhryqrolstyrdlrkgyvypytqkwegelgtldvlsihpgpntvanaiaalte 152
 QY 121 SDKFFINGSNMGGIIGLVAEIAEPDSDLPEPFDLSVKOTHPNFIQLCGAGPELNS 180
 DB 153 sdkffingsnmggiiglvaeiaepdsslepfdsivkothpvnfislcgagpflngs 212
 QY 181 EYLAASVGSMTIIGIDHSLYTGSLWYTPIRREMYEVIIIVRVEINGODIKMCKEYNDK 240
 DB 213 eylvasvgsmtiigidhsltytgslwytptirremyeviiivrveingodikmckeynydk 272
 QY 241 SYVDGSTNLRLPKKVEFAVKSIIKAASSTKEKPPDGFWMGEOLVCOMOAGTTPMNIPEVVS 300
 DB 273 syvdgstnlrlpkkvefaavksikaasstekpdpdgfwlgeqlvcwgaqctpwnlfpvvs 332
 QY 301 LYIMGEVNTQSFRTITLPQOYLRPVEDVATSDDCYKFAISOSTGTWGAIVMEGFYVV 360
 DB 333 lyimgevtngsfrtitlppqylrpvedvatsddcykfaissgstgtwmgavimegfyvv 392
 QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPPVTLDMEDCGYNIPQDESTLMTIAYMAAI 420
 DB 393 fdrarkrigfavsachvhdetrtaavegpfvtldmedcgynipqdestlmtiayymaa 452
 QY 421 CALFMLPICLMWCQWRCLRCLRQOHDDFADDISILK 456
 DB 453 calfmplclmwcwrcrlrclrqghddfaddisilk 488

```
RESULT      3
AAB61334
ID AAB61334 standard; protein; 488 AA.
XX
AC AAB61334;
XX
DT 02-APR-2001 (first entry)
XX
DE Memapsin 2 protein.
XX
KM Memapsin 2; catalyzt; Alzheimer's.
XX
OS Homo sapiens.
XX
PN WO200100663-A2.
XX
PD 04-JAN-2001.
XX
PF 27-JUN-2000; 2000MO-US17661.
XX
PR 28-JUN-1999; 99US-0141363.
PR 30-NOV-1999; 99US-0168060.
PR 25-JAN-2000; 2000US-0177836.
PR 27-JAN-2000; 2000US-0176368.
PR 08-JUN-2000; 2000US-0210292.
XX
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Tang JUN, Lin X, Koelach G;
XX
DR WPI; 2001-102885/11.
XX
PT Purified recombinant catalytically active memapsin 2, used to screen
PT inhibitors of it, which are used to treat and prevent Alzheimer's
PT disease -
XX
PS Claim 2; Page 73-75; 86pp; English.
XX
CC The present invention relates to a purified recombinant
CC catalytically active memapsin 2. The invention may be used for
CC isolating inhibitors which are used to treat or prevent
CC Alzheimer's disease. The invention may also be used to screen
CC for individuals more genetically prone to develop Alzheimer's
CC disease.
XX
SQ Sequence 488 AA;
```

```
Query Match      100.0%; Score 2419; DB 22; Length 488;
Best Local Similarity 100.0%; Pred. No. 4.3e-242;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ETDEEPEEPGRGRSFVEMVDNLGRKSGGYVEMTVGSPPTLNTLVDTGSSNFAVGAAP 60
DB 33 etdeepeepgrgrsfvemdnlgrksgggyvemtvgspptlnlvtgssnfavgaap 92
OY 61 HPEFLRRYQROQLSSTRYRLRKGVYPTYGKWEGLGDLVSIHPGNVTVRANIAITTE 120
DB 93 hpfllrryqrqlsstyrldlrkyvypylcgkwegelgdlvslipgnpvtvranialte 152
OY 121 SDRKFFINSNNEGILGLVAEFAIRPDDSLRPFPSIVKQTHVPMNFSLQCGAGFPNLQS 180
DB 153 sdrkffinsnegilglvaeaetarpddslepfdsivkqthvpmnflscgagfpnlqs 212
OY 181 EYLASVGSMTIIGIDHSLYTGSLMYTPPIRREWYEVIIIVAEINQODLKMDCKEYNYDK 240
DB 213 eylasvgsmtiigldhsllytgslywtpirrewyevliivaeinqdldmckeynydk 272
OY 241 SIYVSGTNNLPLPKKVFPAAKVSKIAASTKRFDPGFMVLGQDLVCWQAGTTPNNIFPVIS 300
DB 273 siyvsgrtnnlplpkkvfpaavksikaastkrfdpdgmvlqgdlvcwqagttppnnifpvis 332
```

```
OY 301 LYLMGEVTNQSPRITTLIPQOYLRPVEDVATSQDDCYKFAISQSGTGYNGAVIMEGEFVV 360
DB 333 lylmgevtntqsprittlipqoylrpvedvatsqddcykfaisqsgtgyngavimegefyvv 392
OY 361 FDRARRIGFAVSACHVHDEPRTAAVEGPFVTLDMDCGVNTPQDESLMTIAVMAAI 420
DB 393 fdrarrigfavsachvhdeprrtaavegpfvtlmdcgvntpqdeslmtiayvmaai 452
OY 421 CALFMLPLCLMWQCMRCLRCLRQOHDHDFADDISLKL 456
DB 453 calfmplclmwqcmrclrclrqhnddfaddislkl 488

RESULT      4
AAY94767
ID AAY94767 standard; Protein; 501 AA.
XX
AC AAY94767;
XX
DT 12-FEB-2001 (first entry)
XX
DE Human beta-secretase amino acid sequence.
XX
KM Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease; human;
KM Down's syndrome; amyloid angiopathy; gene therapy; neuroprotective.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Peptide 1..45
FH /label= putative signal peptide
FT Protein 46..501
FT /label= Beta-secretase
XX
EN WO200058479-A1.
XX
PD 05-OCT-2000.
XX
PF 23-MAR-2000; 2000MO-US07755.
PR 26-MAR-1999; 99US-0277229.
XX
PA (AMGE-) AMGEN INC.
XX
PI Ciltion M, Vassar RJ, Bennett BD;
XX
DR WPI; 2000-594643/56.
DR N-PSDB; AAA28278.
XX
PT Isolated beta-secretase nucleic acids and encoded polypeptides, useful
PT for diagnosis and gene therapy of Alzheimer's disease -
XX
PS Claim 1; Fig 4; 145pp; English.
XX
CC This invention relates to 3 nucleotide sequences encoding beta-secretase
CC proteins. Beta-secretase is an enzyme involved in the production of one
CC of the components of amyloid plaques involved in Alzheimer's disease. The
CC invention includes an expression vector comprising the nucleotide
CC sequence, a host cell comprising the expression vector, and a process for
CC producing the protein through culturing the transformed cells. Also
CC included in the invention are a polypeptide derivative of the
CC beta-secretase protein, a fusion protein comprising beta-secretase fused
CC to a heterologous amino acid sequence, and a method for modulating the
CC levels of beta-secretase polypeptide in a mammal comprising administering
CC the polynucleotide sequence. Beta-secretase exhibits neuroprotecting and
CC neurotropic activity. The beta-secretase nucleotide sequence may be used to
CC map locations of the beta-secretase gene and related genes on chromosomes
CC and as hybridization probes in diagnostic assays to test for the presence
CC of beta-secretase DNA or RNA, such as in Alzheimer's disease, Down's
CC syndrome, and amyloid angiopathy. The nucleotide sequence may also be
CC used as anti-sense inhibitors of beta-secretase expression, in gene
CC therapy of Alzheimer's disease, and for the identification of compounds
```

CC that modulate beta-secretase activity. Antibodies to the beta-secretase
CC protein may be used for in vitro and in vivo diagnostic purposes to
CC detect the presence of beta-secretase polypeptide in a body fluid or cell
CC sample. The present sequence represents the human beta-secretase protein.
CC
XX
SQ Sequence 501 AA;

Query Match 100.0%; Score 2419; DB 21; Length 501;
Best Local Similarity 100.0%; Pred. No. 4.5e-242;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEPEEPGRGSGFVEMVDNLKRGSGGYVEMTVGSPPTLNIIVDTGSSNFAVGAAP 60
DB etdeepeepgrgsgfivemvdnlrkgsggyvemtvgspptlnllydgtssnfavgaap 105
QY 61 HPFLHRYQROLSSYTRDLRKGVVYPYTGKMEGBLGTDLVSIIPHGRNVYRANIAATE 120
DB hpflhryyqrqlsstyrdlrkgyvypycqgkwegelgtvlsviphgpnvtraniaaite 165
QY 121 SDFEFGNSNNEGILGLAVAEIARPDLSLEPFDSLVKQTHVPMLESLQCGAGFPINQS 180
DB sdfefngsnnegilglavaeiarpdlslepfdslvkqthvpmlefsqcgagfpinqs 225
QY 181 EYLASVSGSMITIGTIDHSLYTGSLSMTYPIRREMYEVIIVREINIGODLKMDCKEYNDK 240
DB eylasvsgsmiitigtidhslytgslsmtypirremyeviivreinigodlkmdckeynydk 285
QY 226 evlaavsgsmiitigtidhslytgslytpirremyeviivreinigodlkmdckeynydk 300
DB evlaavsgsmiitigtidhslytgslytpirremyeviivreinigodlkmdckeynydk 345
QY 241 SIVDSGTTNLRPLPKYFPEAAVKSIRKASTSEKFPDGFMLGEOLVCMQAGTPMNIFFPIS 360
DB sivdsgttnlrplpkfyfpeaavksirkaastekfpdgfwlgeqlvcwagttppmniifpvis 405
QY 301 LYLMGEVTNQSFRTITLPPQYLRPVEDVATSDDDCYKFAISQSSGTGYWGAIVMEGYFV 420
DB lylmgevtnqsfrtitlppqylrpvedvatsddcykfaissgstgywgaivmegfyvv 465
QY 406 FDRARRIGFAVSACHVHDEFRTAAVEGPFVLLDMEDCGYNIPQTESTLMTIAYVMAAI 420
DB fdrarrigfavsachvhdefrtaaavgpfvllmedcgyinpqteestlmtiayvmaai 465
QY 421 CALFMLPLCIMVQWRCLRCLROOHDPADDISILK 456
DB 466 calfmplclimvcwrclrcrlqghddfaddisilk 501

RESULT 5
AAB07896
ID AAB07896 standard; Protein; 501 AA.
XX
AC AAB07896;
XX
DT 14-NOV-2000 (first entry)
XX
DE Amino acid sequence of a human beta-secretase enzyme.
XX
KM Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
inhibitor.
XX
OS Homo sapiens.
XX
PN WO200047618-A2.
XX
PD 17-AUG-2000.
XX
PE 10-FEB-2000; 2000MO-US03819.
XX
PR 10-FEB-1999; 99US-0119571.
XX
PR 15-JUN-1999; 99US-0139172.
XX
PA (ELAN-) ELAN PHARM INC.
XX
PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;

PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX
DR WPI: 2000-533011/48.
XX
DR N-PSDB; AAA59550, AAA59551.
XX
PT Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
PS Claim 17; Fig 2A; 121pp; English.

CC The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a human beta-secretase enzyme.

Query Match 100.0%; Score 2419; DB 21; Length 501;
Best Local Similarity 100.0%; Pred. No. 4.5e-242;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEPEEPGRGSGFVEMVDNLKRGSGGYVEMTVGSPPTLNIIVDTGSSNFAVGAAP 60
DB etdeepeepgrgsgfivemvdnlrkgsggyvemtvgspptlnllydgtssnfavgaap 105
QY 61 HPFLHRYQROLSSYTRDLRKGVVYPYTGKMEGBLGTDLVSIIPHGRNVYRANIAATE 120
DB hpflhryyqrqlsstyrdlrkgyvypycqgkwegelgtvlsviphgpnvtraniaaite 165
QY 121 SDFEFGNSNNEGILGLAVAEIARPDLSLEPFDSLVKQTHVPMLESLQCGAGFPINQS 180
DB sdfefngsnnegilglavaeiarpdlslepfdslvkqthvpmlefsqcgagfpinqs 225
QY 181 EYLASVSGSMITIGTIDHSLYTGSLSMTYPIRREMYEVIIVREINIGODLKMDCKEYNDK 240
DB eylasvsgsmiitigtidhslytgslsmtypirremyeviivreinigodlkmdckeynydk 285
QY 226 evlaavsgsmiitigtidhslytgslytpirremyeviivreinigodlkmdckeynydk 300
DB evlaavsgsmiitigtidhslytgslytpirremyeviivreinigodlkmdckeynydk 345
QY 241 SIVDSGTTNLRPLPKYFPEAAVKSIRKASTSEKFPDGFMLGEOLVCMQAGTPMNIFFPIS 360
DB sivdsgttnlrplpkfyfpeaavksirkaastekfpdgfwlgeqlvcwagttppmniifpvis 405
QY 301 LYLMGEVTNQSFRTITLPPQYLRPVEDVATSDDDCYKFAISQSSGTGYWGAIVMEGYFV 420
DB lylmgevtnqsfrtitlppqylrpvedvatsddcykfaissgstgywgaivmegfyvv 465
QY 406 FDRARRIGFAVSACHVHDEFRTAAVEGPFVLLDMEDCGYNIPQTESTLMTIAYVMAAI 420
DB fdrarrigfavsachvhdefrtaaavgpfvllmedcgyinpqteestlmtiayvmaai 465
QY 421 CALFMLPLCIMVQWRCLRCLROOHDPADDISILK 456
DB 466 calfmplclimvcwrclrcrlqghddfaddisilk 501

RESULT 6
AAB6573
ID AAB6573 standard; Protein; 503 AA.
XX
AC AAB6573;
XX
DT 12-APR-2001 (first entry)
XX
DE Human pro-memapsin 2.
XX
KW Human; memapsin 2; neurotrophic; neuroprotective; amyloid precursor protein;

KW APP: memapsin 2 inhibitor; Alzheimer's disease; ss.
XX Homo sapiens.
OS
XX WO200100665-A2.
PN
XX 04-JAN-2001.
PD
XX 27-JUN-2000; 2000WO-US17742.
PF
XX
PR 28-JUN-1999; 99US-0141363.
PR 30-NOV-1999; 99US-0168060.
PR 25-JAN-2000; 2000US-0177836.
PR 27-JAN-2000; 2000US-0178368.
PR 08-JUN-2000; 2000US-0210292.
XX
XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.
PA (UNIT) UNIV ILLINOIS FOUND.
XX
XX Tang JUN, Hong L, Ghosh AK;
PI WPI; 2001-137933/14.
DR
XX
XX Novel memapsin 2 inhibitors which bind to active site of memapsin 2
PT having 2 catalytic aspartic residues and substrate binding cleft, used
PT to treat Alzheimer's disease by blocking amyloid precursor protein
PT cleavage -
XX
XX Example 4; Fig 1; 86pp; English.
PS
XX The present sequence is given in a specification relating to an inhibitor
CC of catalytically active memapsin 2. The inhibitor binds to the memapsin 2
CC active site, which is defined by the presence of two catalytic aspartic
CC residues and a substrate binding cleft. The inhibitor is useful for
CC the treatment and diagnosis of Alzheimer's disease. It is useful in
CC screens for individuals with a genetic predisposition to Alzheimer's
CC disease. The inhibitor is useful as a reagent for specifically binding to
CC memapsin 2 or memapsin 2 analogues and for aiding in memapsin 2
CC isolation, purification and characterisation.
XX
XX Sequence 503 AA:
SO
Query Match 100.0%; Score 2419; DB 22; Length 503;
Best Local Similarity 100.0%; Pred. No. 4.6e-242;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEEPEEPGRGSGFVEMVDNLGRKSGGGYVEMTVGSPQTLNILDVDTGSSNFAYGAAP 60
DB |||||||
QY 48 etdeepeepgrgsfivemvdnlrgksqgyyvemtvgspqtlnlldvtgssnfavgap 107
DB |||||||
QY 61 HPELHRYGROLSSTYRDARKGVYVYTOGKMEGELGDLVSIPHPNVTYRANIAATIE 120
DB |||||||
QY 108 hpeflhrygrqlsstyrdarkgyvyvycqgkwegelgdlvsiphpnvtyraniaaite 167
DB |||||||
QY 121 SDKFFINGSNWEGILGLAVAEIARPDLSLEPFDSLVRKOTHPNLFSLQLCAGFPLNQS 180
DB |||||||
QY 168 sckffingsnwegilglayaeiarpddslpfdsilvkqthvnpnlfslqcgagfplngs 227
DB |||||||
QY 181 EYLAIVGSGSMITIGIDHSLYTGSLWYTPIRREMYEVIIVIVEINGQDLKMDCKEYNTDK 240
DB |||||||
QY 228 evlasvsgsmiigldhsltysglwypirrewyevliivveingqdlkmdckeynydk 287
DB |||||||
QY 241 STVDSGTTNLRPKKVFEEAAVKSIIKAASSTKFPDGFMLGSLQCMQAGTTPMNIFFPVIS 300
DB |||||||
QY 288 slvdsqgttnlrlpkkvfeaaavksikaasstekfpdgfmlgslqcmqagttppmniifpvis 347
QY 301 LYLMGEVYNOSFRITILEQQLRYVEDVATISQDCYKFAISQSSTGTVMGAVIMEGFYVV 360
DB |||||||
QY 348 lylmgevynogfrilileqqrlrvedvatsqdcykfaaisgstgtvmgavimegyvv 407
QY 361 FDRARKRIGFVAVSACHYVDEFTAAVEGPFYTLMEDCGYNIPTDSTLMTIYVMAAI 420
DB |||||||

DB 408 fdrarkrigrfavsachvhelfrtaavegpfvldmedcgyinipqtdeslmtiayvmaai 467
QY 421 CALFMLPLCLMWCOMRCLRCLROQHDDFPADISTILK 456
DB |||||||
DB 468 calfmlplclmwcomrclrlrqhddfadadistilk 503
RESULT 7
AAB61335
ID AAB61335 standard; protein; 503 AA.
XX
AC AAB61335;
XX
DT 02-APR-2001 (first entry)
XX
DE T7 promoter and vector sequence.
XX
KW Memapsin 2; catalyst; Alzheimer's.
XX
XX Homo sapiens.
OS Synthetic.
XX
PN WO200100663-A2.
XX
XX 04-JAN-2001.
PD
XX 27-JUN-2000; 2000WO-US17661.
PF
XX 28-JUN-1999; 99US-0141363.
PR 30-NOV-1999; 99US-0168060.
PR 25-JAN-2000; 2000US-0177836.
PR 27-JAN-2000; 2000US-0178368.
PR 08-JUN-2000; 2000US-0210292.
XX
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
PI Tang JUN, Ian X, Koelsch G;
XX
XX WPI; 2001-102885/11.
DR
XX
XX Purified recombinant catalytically active memapsin 2, used to screen
PT inhibitors of it, which are used to treat and prevent Alzheimer's
PT disease -
XX
PS Disclosure; Fig 1; 86pp; English.
XX
XX The present invention relates to a purified recombinant
CC catalytically active memapsin 2. The invention may be used for
CC isolating inhibitors which are used to treat or prevent
CC Alzheimer's disease. The invention may also be used to screen
CC for individuals more genetically prone to develop Alzheimer's
CC disease.
XX
XX Sequence 503 AA:
SO
Query Match 100.0%; Score 2419; DB 22; Length 503;
Best Local Similarity 100.0%; Pred. No. 4.6e-242;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEEPEEPGRGSGFVEMVDNLGRKSGGGYVEMTVGSPQTLNILDVDTGSSNFAYGAAP 60
DB |||||||
QY 48 etdeepeepgrgsfivemvdnlrgksqgyyvemtvgspqtlnlldvtgssnfavgap 107
DB |||||||
QY 61 HPELHRYGROLSSTYRDARKGVYVYTOGKMEGELGDLVSIPHPNVTYRANIAATIE 120
DB |||||||
QY 108 hpeflhrygrqlsstyrdarkgyvyvycqgkwegelgdlvsiphpnvtyraniaaite 167
DB |||||||
QY 121 SDKFFINGSNWEGILGLAVAEIARPDLSLEPFDSLVRKOTHPNLFSLQLCAGFPLNQS 180
DB |||||||
QY 168 sckffingsnwegilglayaeiarpddslpfdsilvkqthvnpnlfslqcgagfplngs 227
QY 181 EYLAIVGSGSMITIGIDHSLYTGSLWYTPIRREMYEVIIVIVEINGQDLKMDCKEYNTDK 240
DB |||||||

```

|||||
Db 228 EVLASVGSMTIGIDHSLYGSILWYPIRREWEYEVILVIVEINGDLKMDCKEYNDK 287
QY 241 SIYDSGTNLRPKKVEAAVKSIAASTTEKFPDGFNLGSLYCWQAGTTPNNIPVIS 300
Db 248 SIVDSGTNLRPKKVEAAVKSIAASTTEKFPDGFNLGSLYCWQAGTTPNNIPVIS 347
QY 301 LYIMGEVYNOSFRTILPQOYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYV 360
Db 348 IYIMGEVYNOSFRTILPQOYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYV 407
QY 361 FDRARRKRGFAVSACHVDEFTAAVEGPEVTLDMEDCGYVNPOTDESTLMTIAYVMAAI 420
Db 408 fdrarkrigrfavsachvdeftaaavgpfvtlmedcgyinpqtdestlmtlayvmaai 467
QY 421 CALFMLPCLMWCOMRCRLRQOHDDFADDISILK 456
Db 468 calfmplclmwcomrcrlrqohddaddisilk 503

RESULT 8
AAE10629 standard; Protein: 501 AA.
AAE10629;
AAE10629;
03-AUG-2000 (first entry)
Human aspartyl protease 2 (a) (Asp2) amino acid sequence.
Aspartyl protease; aspartase; amyloid precursor protein; APP; Asp 2;
Alzheimer's disease; beta secretase site.
Homo sapiens.
WO200017369-A2.
30-MAR-2000.
23-SEP-1999; 99WO-US20881.
24-SEP-1998; 98US-0101594.
(PHAA) PHARMACIA & UPJOHN CO.
Gurney ME, Bienkowski MJ, Heinrikson RL, Parodi LA, Van R;
WPI: 2000-303209/26.
N-PSDB; AAA15662.
New enzyme designated human aspartase useful in research into
Alzheimer's Disease is capable of cleaving amyloid protein precursor at
the beta secretase site to produce amyloid beta peptide.
Claim 48; Fig 2; 183pp; English.
This sequence represents the human aspartyl protease 2 (Asp2) amino acid
sequence. The invention relates to a protease (e.g. Asp2) capable of
cleaving the beta secretase site of amyloid precursor protein (APP). The
sequence contains a sequence encoding the amino acid sequence DTG and a
sequence encoding DSG or DTG separated by 100-300 amino acids. When
mutated the APP gene causes an autosomal dominant form of Alzheimer's
disease. APP localises to the cell surface membrane and have a single
C-terminal transmembrane domain. Proteolytic processing of APP produces
the amyloid beta protein, which is possibly very important in Alzheimer's
disease. The invention includes a nucleotide sequence encoding the
protease, a vector containing the nucleotide sequence, and a cell line
comprising the vector. Methods for screening for inhibitors of beta
secretase activity are also given in the invention. The human aspartase
protein and nucleotide sequences and the methods for identifying
inhibitors of the protease, are useful in the treatment of and research
in to Alzheimer's disease.

```

```

SQ Sequence 501 AA:
Query Match 99.8%; Score 2414; DB 21; Length 501;
Best Local Similarity 99.8%; Pred. No. 1,5e-241;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ETDEEPEERGRSGFVEMVNDLKGSGGGYVEMTQSPROTINILVDPGSSNFAYGAAP 60
Db 46 etdeepeepgrsgfvevmdnlkgsggyvemtvsppqtnllvdsgrsfaygaap 105
QY 61 HPFLHRYQROLSTSYRDLKRGYVPTQGMGEGELGTDLVSIHPGPNVTNRANIAITE 120
Db 106 hpflhryqrolstsyrdlkrqyvpvtqgmegeltldvsihpgpnvtvraniaite 165
QY 121 SDRKFFGSMWEGILGLAYAEIARPDLSLEPFDDSLVKQKHVNNLESLQCGAFPLNOS 180
Db 166 sdrkffgsmwegilglayaeiarpdlslepfdslvkqkhvnnleslncgafplngs 225
QY 181 EVLASVGSMTIGIDHSLYGSILWYPIRREWEYEVILVIVEINGDLKMDCKEYNDK 240
Db 226 EVLASVGSMTIGIDHSLYGSILWYPIRREWEYEVILVIVEINGDLKMDCKEYNDK 285
QY 241 SIYDSGTNLRPKKVEAAVKSIAASTTEKFPDGFNLGSLYCWQAGTTPNNIPVIS 300
Db 286 SIVDSGTNLRPKKVEAAVKSIAASTTEKFPDGFNLGSLYCWQAGTTPNNIPVIS 345
QY 301 LYIMGEVYNOSFRTILPQOYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYV 360
Db 346 IYIMGEVYNOSFRTILPQOYLRPVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYV 405
QY 361 FDRARRKRGFAVSACHVDEFTAAVEGPEVTLDMEDCGYVNPOTDESTLMTIAYVMAAI 420
Db 406 fdrarkrigrfavsachvdeftaaavgpfvtlmedcgyinpqtdestlmtlayvmaai 465
QY 421 CALFMLPCLMWCOMRCRLRQOHDDFADDISILK 456
Db 466 calfmplclmwcomrcrlrqohddaddisilk 501

RESULT 9
AAE10629 standard; Protein: 501 AA.
AAE10629;
AAE10629;
10-DEC-2001 (first entry)
Human aspartyl protease 2(a) [hu-Asp2(a)] protein.
Human; aspartyl protease 2(a); Asp2(a); amyloid precursor protein; APP;
Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
amyloid plaque; neuronal loss; proteolytic; neurotrophic; neuroprotective;
chromosome 11q23.3-24.1.
Homo sapiens.
Key Location/Qualifiers
Peptide 1..21
Peptide /label= Signal_peptide
Peptide 22..45
Peptide /label= Asp_2a_prepropeptide
Peptide 46..57
Peptide /label= Asp_2a_propeptide
Protein 58..501
Protein /label= Mature_human_Asp_2a_protein
Region 420..454
Region /label= Alpha_helical_spacer_region
Domain 455..477
Domain /label= Transmembrane_domain
Domain 478..501
Domain /label= Cytoplasmic_domain

```


PN GB235767-A.
 XX
 PD 04-JUL-2001.
 XX
 PF 22-SEP-2000; 2000GB-0023315.
 XX
 PR 23-SEP-1999; 99US-0155493.
 PR 23-SEP-1999; 99US-0404133.
 PR 23-SEP-1999; 99WO-US20881.
 PR 13-OCT-1999; 99US-0416901.
 PR 06-DEC-1999; 99US-0169232.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 PI Bienkowski MJ, Gurney M;
 DR MPI; 2001-444208/48.
 DR N-PSDB; AAD17865.
 XX
 XX
 PT Polypeptide comprising fragments of human aspartyl protease with
 PT amyloid precursor protein processing activity and alpha-secretase
 PT activity, for identifying modulators useful in treating Alzheimer's
 PT disease -
 XX
 PS Example 2; Fig 2; 187pp; English.
 XX
 CC The patent discloses human aspartyl protease 1 (hu-Asp1) or modified
 CC Asp1 proteins which lack transmembrane domain or amino terminal
 CC domain or cytoplasmic domain and retains alpha-secretase activity
 CC and amyloid protein precursor (APP) processing activity. The proteins
 CC of the invention are useful for assaying hu-Asp1 alpha-secretase
 CC activity, which in turn is useful for identifying modulators of
 CC hu-Asp1 alpha-secretase activity, where modulators that increase
 CC hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's
 CC disease (AD) which causes progressive dementia with consequent
 CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
 CC neuronal loss. Hu-Asp1 protease substrate is useful for assaying
 CC hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with
 CC the substrate under acidic conditions and determining the level of
 CC hu-Asp1 proteolytic activity. The present sequence is long form of
 CC human Asp2 protein, designated as Asp2(a). Asp2 gene is localised
 CC on chromosome 11q23.3-24.1.
 CC
 SQ Sequence 501 AA:
 Query Match 99.8%; Score 2414; DB 22; Length 501;
 Best Local Similarity 99.8%; Pred. No. 1.5e-241;
 Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 EFDDEPEEPGRGSGFVEMVDNRKSGGQYVEMTVGSPPTQTLNLTVDTGSSNFAVGAAP 60
 Db 46 etdeepeepgrgsgfivemvdnlrtksgqgyvemtvgspptqlnlvtgssnfavgap 105
 Oy 61 HPELHRYTORQLSSTYRDLRKGVYVYTOGKWEGLGDIVSIPHGPRVTVRANAATTE 120
 Db 106 hpehlrhytorglsstyrldlrkgyvyvtyogkwegelgdivsiphgprvtrvanaatle 165
 Oy 121 SDFKFNCSNMWGIIGLAVYAEIARPDLSLEPPFDSLVTQTHVNLFSLQLGAGFPLNQS 180
 Db 166 sdfkfnsgnmwgiiglavyaetiarpdlsleppfDSLVTQTHVNLFSLQLGAGFPLNQS 225
 Oy 181 EYLAIVGSGSMITGIDHLYGSLMYTPYRREKRYEVIIVAEINSGDLKMDCKEYVNDK 240
 Db 226 eylavsgsgsmitgldhlygslmytpyrrekyeviivaeinSGDLKMDCKEYVNDK 285
 Oy 241 SIVDSGTTLNLPRKVFEEAAVKSIAASSTEFKPDGFNLGEOLVCMQMGCTPWNINFPITS 300
 Db 286 sivdsgttnlprkvfEEAAVKSIAASSTEFKPDGFNLGEOLVCMQMGCTPWNINFPITS 345
 Oy 301 LYLMGEVYNOSFRITILFQOYLRFVEDVATSQDDCYKFAISQSSGTGVMAVINEGEFVV 360
 Db 346 lylmgevynosfrtilfpgylrfvedvatsqddcykfaisqssgtgvmavinegefvyv 405

Oy 361 FDRARRIGFANYSACHVHDEFRTRAVEGPFVTLNMDGCGYNIPQDESTLMTIAYVMAI 420
 Db 406 fdrarrigfavsachvhdefrtRAVEGPFVTLNMDGCGYNIPQDESTLMTIAYVMAI 465
 Oy 421 CALFMDPLCIAMVCOMRCRLRQOHDDFADDISLTK 456
 Db 466 calfmplcmvcomrcrlrqlrqohddadddisltk 501
 RESULT 10
 AAE06859
 ID AAE06859 standard; Protein; 501 AA.
 XX
 AC AAE06859;
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE Human aspartyl protease 2a (Hu-Asp2a) protein.
 XX
 KM Human; aspartyl protease 2a; Asp 2a; beta-amyloid precursor protein; APP;
 KM beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
 KM neurofibrillary tangle; neuronal loss; amyloid-beta peptide; neurotropic;
 KM neuroprotective; antisense therapy; gene therapy;
 KM chromosome 11q23.3-24.1.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Peptide 1..21
 FH /label= Signal_peptide
 FH Protein 22..501
 FH Region /note= "Mature human aspartyl protease 2a (Hu-Asp2a)"
 FH Domain 420..454
 FH /note= "Alpha helical spacer region"
 FH Domain 455..477
 FH /label= Transmembrane_domain
 FH Domain 478..501
 FH /label= Cytoplasmic_domain
 PN WO200150829-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 09-MAY-2001; 2001WO-IB00799.
 XX
 PR 09-MAY-2001; 2001WO-IB00799.
 XX
 PA (BIEN/) BIENKOWSKI M J.
 PA (GURN/) GURNEY M E.
 PA (HEIN/) HEINRIKSON R L.
 PA (PARO/) PARODI L A.
 PA (YANR/) YAN R.
 XX
 PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;
 DR MPI; 2001-483072/52.
 DR N-PSDB; AAD13021.
 XX
 PT Novel purified polypeptide comprising fragment of mammalian aspartyl
 PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 PT activity -
 XX
 PS Claim 49; Fig 2; 185pp; English.
 XX
 CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
 CC precursor protein (APP) isoforms and their corresponding DNA molecules.
 CC Human aspartyl proteases can act as beta-secretase proteases useful for
 CC treating Alzheimer's disease. APP isoforms are useful for identifying
 CC modulators of amyloid-beta peptide production, for use in designing
 CC therapeutics for the treatment and prevention of Alzheimer's disease,
 CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis

CC and neuronal loss. APP isoforms are also used in methods for identifying
 CC inhibitors and modulators of human Asp2 activity. The invention relates
 CC to a method for identifying agents that modulate the activity of human
 CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
 CC as a means to screen in cellular assays for the inhibitors of beta- and
 CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
 CC polymerase chain reactions (PCR). The probes are useful for detecting
 CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
 CC blots. The present sequence is human aspartyl protease 2 (Hu-Asp2), a
 CC 'long' form designated as (Hu-Asp2a). Hu-Asp 2 gene is localised on
 CC chromosome 11q23.3-24.1.

XX Sequence 501 AA:

Query Match 99.8%; Score 2414; DB 22; Length 501;
 Best Local Similarity 99.8%; Pred. No. 1.5e-241;
 Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEEPEEPRGRSGFVEMVDNLKSGSGGYVEMTVGSPQTLNLIIVDGSSENFAYGAAP 60
 |||||||
 Db 46 etdeepeepgrsgfivemvndlrksgsggyvemtvgspqtlnlivdgsseinfaygaap 105
 |||||||
 QY 61 HPELHRYGROLSTSYRDLRKGVYVPTGKMGELGTDLVSIIPHGPNTVTRANIAAITE 120
 |||||||
 Db 106 hpeflhryyrolsstyrdlrkgyvpyvtgkwegelgtldvsiiphgpnvtraniaaite 165
 |||||||
 QY 121 SDKFINGSNWEGIGLAVAEIARPDLSIEPPDLSIKQTHVNFSLDLCAGAPFLNOS 180
 |||||||
 Db 166 sdkffingsnwegilglavaeiarpdlsiepfdsivkqthvnlfsllcgaqfplnos 225
 |||||||
 QY 181 EYIASVSGMIIGIDHSIYTGSLWYTPIRREYEVIIIVAREINQDLMCKEKNYDK 240
 |||||||
 Db 226 evlasvsgsmliigidhslytgslywtpirreyeviiivareinqgdlnmckekeynk 285
 |||||||
 QY 241 SYDSGTTNLRLPKRYFEAAVKSIRKASSTKRPDGFMLGEOIIVCMQAGTPWNIPVTS 300
 |||||||
 Db 286 sydsgstnlrlpkryfeaaavksirkaasstekrpdgfwlgeqivcwaqgtlpwnlfpvts 345
 |||||||
 QY 301 LYTMGEVYNOSFRTITLLPQOYLRFVSDVATSQDDCYKFAISQSSCTGTVMGAVIMEGFYV 360
 |||||||
 Db 346 lytmgevtnsfritllpqoylrfvsvdatSQDDCYKFAISGSSCTGTVMGAVIMEGFYV 405
 |||||||
 QY 361 FDMARRIGAVASACHVHDEFRTAAVEGPFVTMDMDCGYNIPQTESTLMTIAYMAAI 420
 |||||||
 Db 406 fdmarrigavasachvndeifrtaaavgpfvtldmedcgyndipqtdestlmtlaymaai 465
 |||||||
 QY 421 CALFMLPLCLMVCWMRCLRCLROQHDDFADDISLILK 456
 |||||||
 Db 466 calfmplclmvcwrcrlclrqghddfaddislilk 501
 |||||||

RESULT 11

AAU06603 ID AAU06603 standard; Protein; 501 AA.

XX AAU06603;

XX 24-OCT-2001 (first entry)

XX Human Aspartyl protease 2(a), Asp2(a).

XX Human: Aspartyl protease; Asp2(a); beta-secretase; neurotrophic;

XX neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;

XX amyloid-beta; Abeta.

XX Homo sapiens.

XX Key

FT Peptide 1..21 Location/Qualifiers
 FT Peptide /label= signal peptide
 FT Peptide 22..45
 FT Peptide /label= pre-pro-peptide

FT Peptide 46..57
 FT /label= Pro-peptide
 FT Protein 57..501
 FT /label= Mature-Asp2(a)
 FT Region 420..454
 FT /label= Alpha-helical_spacer_region
 FT Domain 455..477
 FT /label= Transmembrane_domain
 FT Domain 478..501
 FT /label= cytoplasmic_domain

PN W0200149098-A2.

PD 12-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00798.

XX 09-MAY-2001; 2001WO-IB00798.

PA (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

PA (YANR/) YAN R.

PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;

XX WPI: 2001-502549/55.

DR N-PSDB; AAS11517.

PT Novel purified polypeptide comprising fragment of mammalian aspartyl

PT protease 2, lacking Asp2 transmembrane domain and retaining beta

PT secretase activity of Asp2 useful for identifying inhibitors of Asp2

PT activity -

PS Claim 49; Fig 2; 185pp; English.

XX The invention relates to a purified polypeptide comprising a fragment of

CC mammalian aspartyl protease (Asp2) protein which lacks the Asp2

CC transmembrane domain and the Asp2 protein, and where the polypeptide and

CC the fragment retain the beta-secretase activity of the mammalian Asp2

CC protein. The invention also details polynucleotides for the Asp

CC proteins and vectors expressing them, and a polypeptide (isoform of

CC amyloid protein precursor (APP) comprising the amino acid sequence of an

CC APP or its fragment containing an APP cleavage site recognizable by a

CC mammalian beta-secretase, and further comprising two lysine residues at

CC the carboxyl terminus of the amino acid sequence are methods of identifying

CC APP fragment. Also included in the invention are methods of identifying

CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are

CC useful for treating Alzheimer's disease. APP is useful in methods for

CC identifying inhibitors or modulators of human Asp2 activity and

CC amyloid-beta (Abeta) peptide production. APP is also useful in designing

CC therapeutics for the treatment or prevention of Alzheimer's disease.

CC APP comprising the APP-SW-beta-secretase peptide sequence (NDA), which

CC is associated with increased levels of Abeta processing is useful in

CC assays relating the Alzheimer's research. The expression vector is useful

CC for recombinantly expressing APP. Nucleic acids that hybridise to

CC Asp oligonucleotides are useful as probes or primers. The probes are

CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in

CC Northern and Southern blots. The present sequence is human Asp2(a).

XX Sequence 501 AA:

Query Match 99.8%; Score 2414; DB 22; Length 501;
 Best Local Similarity 99.8%; Pred. No. 1.5e-241;
 Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEEPEEPRGRSGFVEMVDNLKSGSGGYVEMTVGSPQTLNLIIVDGSSENFAYGAAP 60
 |||||||
 Db 46 etdeepeepgrsgfivemvndlrksgsggyvemtvgspqtlnlivdgsseinfaygaap 105
 |||||||
 QY 61 HPELHRYGROLSTSYRDLRKGVYVPTGKMGELGTDLVSIIPHGPNTVTRANIAAITE 120
 |||||||

Db 106 hpfllhyrqlststydrlrkgyvpytgkgwgelgldslviphgpnvtraniaalte 165
QY 121 SDFEFGNSNMGIIIGLAFAEIAARPDDSEFPFDSLVKQTHVNLFSLOLCGAGFPLNQS 180
Db 166 sckffngsnwgeglglayaelarppdslepfidslvkqthvpnlfsllcgagfplnqs 225
QY 181 EVLASVGSMIIGIDHSILYTGSLWYTPIRREMYEVIIVREINGODLKMCKEYNDK 240
Db 226 evlasvgsmligldhslaytgslytprlrewyevllvireingqdlkmdckeynydk 285
QY 241 SIYDSGTTNLRPKKVFEEAAYKSIKAASSTEFKPDGFWLGEOLVCMQAGTTPANNIFPVIS 300
Db 286 slvdsqgtlnlrpkkvfaeavksikaasstekfpdgfwlgeqlvcwgagctpwnlfpvls 345
QY 301 LYLMGEVTVNOSFRITILPQOYLRPVEDVATSDDCYKFAISQSGTGYVMGAVIMEGFYV 360
Db 346 lymgevtngsfrltllpqyrlrpedvatsqddcykfalsqsgtvmgavimegfyyv 405
QY 361 FDRARRKRGFAVSACHVDEFRRTAAVEGPFVTLDMEDCGYNIPOTDSTLMTIAYVMAAI 420
Db 406 fdrarrkrigfavsachvdefrtaavegpfvclmedcgynipqtdestlmtlayymaa 465
QY 421 CALFMLPLCIWCOMRCRLRQOHDDFADDSILK 466
Db 466 calfmplclmwgcwrcrlrcrlqghddfaddslilk 501

RESULT 12

AAU07202 standard; Protein; 501 AA.

AAU07202:

24-OCT-2001 (first entry)

Human aspartyl protease 2a (Asp-2a).

Human: aspartyl protease 1; Asp-1; nootropic; neuroprotective;
Kw aspartyl protease 2; Asp2; amyloid protein precursor; App;
Kw beta-secretase; Alzheimer's disease.

OS Homo sapiens.

Key Location/Qualifiers
peptide 1..21

Misc_feature 22..45 /note= "Signal peptide"

Misc_feature 46..57 /note= "Pre-propeptide"

Protein 58..501 /note= "Propeptide"

Region 420..454 /note= "Mature Aspartyl protease-2a"

Domain 455-477 /note= "Alpha helical spacer region"

Domain 478..501 /note= "Transmembrane domain"

Domain /note= "Cytoplasmic domain"

WO200149097-A2.

12-JUL-2001.

09-MAY-2001; 2001WO-IB00797.

09-MAY-2001; 2001WO-IB00797.

(BIEN/) BIENKOWSKI M J.

(GURNEY/) GURNEY M E.

(HEIN/) HEINRIKSON R L.

(PARO/) PARODI L A.

(YANR/) YAN R.

XX
PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX
XX WPI: 2001-502548/55.
DR N-PSDB; AASL1702.
XX
PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
XX
PS Claim 49; Fig 2; 185pp; English.
XX
CC The invention relates to a novel purified polypeptide comprising a
CC fragment of mammalian aspartyl protease 2 (Asp2) protein which lacks the
CC Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide
CC and the fragment retain the beta-secretase activity of the mammalian Asp2
CC protein. Also included is an isoform of amyloid protein precursor (APP)
CC comprising the amino acid sequence of a APP or its fragment containing
CC an APP cleavage site recognizable by a mammalian beta-secretase, and
CC further comprising two lysine residues at the carboxyl terminus of the
CC amino acid sequence of the mammalian APP or APP fragment. The
CC polypeptides are used for assaying for modulators of beta-secretase
CC activity; identifying agents that inhibit the APP processing activity
CC of human Asp2 aspartyl protease (Hu-Asp2); identifying agents that
CC modulate the activity of Asp2; and for reducing cellular production of
CC amyloid beta (Abeta) from APP. Agents identified by the above methods
CC are useful for treating Alzheimer's disease; and for identifying
CC modulators of amyloid-beta (Abeta) peptide production, for use in
CC designing therapeutics for the treatment or prevention of Alzheimer's
CC disease. Probes and primers derived from Asp nucleic acid sequences
CC are useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC Northern and Southern blots. The present sequence represents the
CC amino acid sequence of human Asp-2a used in the methods of the invention.
XX
SQ Sequence 501 AA;

Query Match 99.88; Score 2414; DB 22; Length 501;
Best Local Similarity 99.88; Pred. No. 1,5e-241;

Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEPEEPGRGSEFVMDLNRKSGGYVEMTVGSSPOTINIIVDSSNFANAGAP 60
Db 46 etdepeepgrgsfvmvnlrkgsggyvemtvgsspqtlnllvdygssnfavgaap 105
QY 61 HPELHRYQRLSTYDLRKGYVVPYTGKGEGELGTDLVSIIPHGPNTVRANIAITE 120
Db 106 hpfllhyrqlststydrlrkgyvpytgkgwgelgldslviphgpnvtraniaalte 165
QY 121 SDFEFGNSNMGIIIGLAFAEIAARPDDSEFPFDSLVKQTHVNLFSLOLCGAGFPLNQS 180
Db 166 sckffngsnwgeglglayaelarppdslepfidslvkqthvpnlfsllcgagfplnqs 225
QY 181 EVLASVGSMIIGIDHSILYTGSLWYTPIRREMYEVIIVREINGODLKMCKEYNDK 240
Db 226 evlasvgsmligldhslaytgslytprlrewyevllvireingqdlkmdckeynydk 285
QY 241 SIYDSGTTNLRPKKVFEEAAYKSIKAASSTEFKPDGFWLGEOLVCMQAGTTPANNIFPVIS 300
Db 286 slvdsqgtlnlrpkkvfaeavksikaasstekfpdgfwlgeqlvcwgagctpwnlfpvls 345
QY 301 LYLMGEVTVNOSFRITILPQOYLRPVEDVATSDDCYKFAISQSGTGYVMGAVIMEGFYV 360
Db 346 lymgevtngsfrltllpqyrlrpedvatsqddcykfalsqsgtvmgavimegfyyv 405
QY 361 FDRARRKRGFAVSACHVDEFRRTAAVEGPFVTLDMEDCGYNIPOTDSTLMTIAYVMAAI 420
Db 406 fdrarrkrigfavsachvdefrtaavegpfvclmedcgynipqtdestlmtlayymaa 465
QY 421 CALFMLPLCIWCOMRCRLRQOHDDFADDSILK 466
Db 466 calfmplclmwgcwrcrlrcrlqghddfaddslilk 501

RESULT 13
AAE02581
ID AAE02581 standard; Protein: 501 AA.
XX
AC AAE02581;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human aspartyl protease 2a (asp 2a).
XX
KW Human: alpha-secretase; amyloid precursor protein; APP; therapy;
KM Alzheimer's disease; antialzheimer's; aspartyl protease 2a; Asp 2a;
KM beta-secretase; chromosome 11q23.3-24.1.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT Peptide
FT Peptide
FT Peptide
FT Peptide
FT Protein
FT Active-site
FT Active-site
FT Active-site
FT Region
FT Region
FT Domain
FT Domain
FT Domain
FT Region
XX
XX WO200123533-A2.
XX
XX PD 05-APR-2001.
XX
XX PE 22-SEP-2000; 2000WO-US26080.
XX
XX PR 23-SEP-1999; 99US-0155493.
XX PR 23-SEP-1999; 99WO-US20881.
XX PR 13-OCT-1999; 99US-0416901.
XX PR 06-DEC-1999; 99US-0169232.
XX
XX PA (PHAA) PHARMACIA & UPJOHN CO.
XX
XX PI Gurney M, Bienkowski MJ;
XX
XX DR WPI: 2001-290516/30.
XX
XX DR N-PSDB; AAD06739.
XX
XX PT Enzymes that cleave the alpha-secretase site of the amyloid precursor
XX PT protein, useful for the treatment of Alzheimer's disease -
XX
XX PS Example 2; Fig 2; 189pp; English.
XX
XX CC The present invention relates to enzymes for cleaving the alpha-
XX CC secretase site of the amyloid precursor protein (APP) and methods of
XX CC identifying those enzymes. The methods may be used to identify enzymes
XX CC that may be used to cleave the alpha-secretase cleavage site of the APP
XX CC protein. The enzymes may be used to treat or modulate the progress of
XX CC Alzheimer's disease. The present sequence is human aspartyl protease 2a
XX CC (Asp 2a). Asp 2a has beta-secretase protease activity. Asp2 gene
XX CC is located on chromosome 11q23.3-24.1.
XX
XX SQ Sequence 501 AA;

Query Match 99.8%; Score 2414; DB 22; Length 501;
Best Local Similarity 99.8%; Pred. No 1,5e-241;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ETDEEPEEPGRGSGFVEMVDNLGKSGGYYVEMTVGSPPTLNLVDTGSSNFAVGAAP 60
DB 46 etdeepeepgrirgsifewmdnlrgksgggyvemtvgspptclnlivdsgsnfavgap 105
QY 61 HPELHRYQRLSSTTRDLRKGYVYPTTGKMEGELGTDLVSPHGPNTVRAANIAITE 120
DB 106 hpeLhryqRlSstTRdLrKgyvYpTtGkMeGeLgTdLvSpHgpNtVraNiaIte 165
QY 121 SDFEFGSNWEGILGAYAEIARPDLSLEPPDSLVKQTHVNFSLQCGAGFPLNS 180
DB 166 sdfEfgsnWegIlGAYaeIARpDLSlePPdSLvKqThvNfSLqCGaGFpLns 225
QY 181 EVLASVGSMLIGIDHSLYTGSLSWTPPIREMYEVITYRVEINGODLKMCKEYNYDK 240
DB 226 evlasvgsmliGidHsLyTgSlsWtPPIrEmYEvItYrVeInGODlKMcKeYnYdk 285
QY 241 STVDSGTTNRLPKKVFEEAVKSIKAASSTKFPDGFMLGEOLVWQAGTTPWNIFPVIS 300
DB 286 stvdsGttnrLpKkVfEEaVksIkAasStKfPdGfMlGeOlVwQAGtTPwnIfPvis 345
QY 301 LYLMGEVTNQSFRITILPQOYLRPVEDVATSDDCYKFAISQSTGTYMKAVIMEGFYV 360
DB 346 lylmgevtnqsfrItIlPqOyLrpVeDvAtsDdcYkFaiSqStGtyMkAVImEGfYv 405
QY 361 FDRARRKRGFAVSACHVDEFTAAVEGPFVTLDMECGYNIPTDDESTMTIAYVMAAT 420
DB 406 fdrArkrRGfAVsACHvDEfTAAveGPFvTLdMEcGYnIPTdDEsTMTIayVmaAt 465
QY 421 CALFMLPLCLMWQWRCRLRQOHDDFADDISLTK 456
DB 466 calfmlpLclmwqWrcLrQohDdfAdDislTk 501
RESULT 14
AAW59807
ID AAW59807 standard; Protein: 501 AA.
XX
XX AC AAW59807;
XX
XX DT 26-OCT-1998 (first entry)
XX
XX DE Amino acid sequence of human ASP2 (aspartic protease 2).
XX
XX KW Human: ASP2; aspartic protease 2; agonist; antagonist; immunospecific;
XX KW antibody; inhibition; Alzheimer's disease; cancer; proteinase;
XX KW prohormone processing.
XX
XX OS Homo sapiens.
XX
XX OS EP855444-A2.
XX
XX PN EP855444-A2.
XX
XX PD 29-JUL-1998.
XX
XX PF 27-JAN-1998; 98EP-0300573.
XX
XX PR 28-JAN-1997; 97GB-0001684.
XX
XX PA (SMIK) SMITHKLINE BEECHAM CORP.
XX PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
XX PI Chapman CG, Murphy K, Powell DJ, Smith TS;
XX
XX DR WPI: 1998-389809/34.
XX
XX DR N-PSDB; AAVA1696.
XX
XX PT New nucleic acid encoding human aspartic protease 2 - used to treat,
XX PT prevent and diagnose e.g. Alzheimer's disease, cancer and prohormone

PT processing
XX
PS Claim 1; Page 7; 26pp; English.
XX
CC This is the amino acid sequence of the human ASP2 (aspartic protease
CC family), used in the method of the invention. Agonists and
CC antagonists for ASP2 immunospecific antibodies are used to treat
CC conditions requiring increased or decreased activity or expression of
CC ASP2 respectively. ASP2 is used to treat and diagnose e.g.,
CC Alzheimer's disease, cancer and pro-hormone processing and ASP2 or a
CC fragment can be used to induce an immune response against the above
CC conditions.
XX
SQ Sequence 501 AA;

Query Match 99.8%; Score 2413; DB 19; Length 501;
Best Local Similarity 99.8%; Pred. No. 1,9e-241;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ETDEPEEPGRGSGFVEMVDNLKSGSGGYVEMTVGSPPTQNLIVDTGSSNFAVGAAP 60
DB etdepeepgrgsgfvevmdnlrgksqgyvemtvgspptqnlivdtgssnfavgaap 105
QY 61 HPFLHRYRQRLSTYRDLRKGYVVPYTOGKWESELTDLVSIHPGPNVTVRANIAITE 120
DB hpfllhryrqrlstydrlrkgyvvpptogkwezeltdlvsihpgpnvtvraniaite 165
QY 121 SDRFFINGSMWEGILGLAVAEIARPDLSLEPFDSLVKOTHPVNLPSLQCGAGFPLNOS 180
DB sdrffingsmwegilglavaeiarpdslepfdslvkothpvnlpsslqcgagfplnds 225
QY 166 Sdkffingsmwegilglavaelarppdslepfdsivkqthpnlsfqlcgagfplnds 225
DB 166 Sdkffingsmwegilglavaelarppdslepfdsivkqthpnlsfqlcgagfplnds 225
QY 181 EVLASVGSMTIGIDHSILYTGSLMTPTPIRREYVYVIVRVEINQDLMCKEYNDK 240
DB 181 EVLASVGSMTIGIDHSILYTGSLMTPTPIRREYVYVIVRVEINQDLMCKEYNDK 240
QY 226 EVLASVGSMTIGIDHSILYTGSLMTPTPIRREYVYVIVRVEINQDLMCKEYNDK 285
DB 226 EVLASVGSMTIGIDHSILYTGSLMTPTPIRREYVYVIVRVEINQDLMCKEYNDK 285
QY 241 SIVDSGTTNLRPKKVEFAVKSIRKASTKEKPPDFGFWGEOIYVQOACTPTPNTPIVPS 300
DB 241 SIVDSGTTNLRPKKVEFAVKSIRKASTKEKPPDFGFWGEOIYVQOACTPTPNTPIVPS 300
QY 286 SIVDSGTTNLRPKKVEFAVKSIRKASTKEKPPDFGFWGEOIYVQOACTPTPNTPIVPS 345
DB 286 SIVDSGTTNLRPKKVEFAVKSIRKASTKEKPPDFGFWGEOIYVQOACTPTPNTPIVPS 345
QY 301 LYLMGEVTNOSRITILPQOYLRPVEDVATSDDCYKFAISQSTGTGAVIMEGFYV 360
DB 301 LYLMGEVTNOSRITILPQOYLRPVEDVATSDDCYKFAISQSTGTGAVIMEGFYV 360
QY 346 LYLMGEVTNOSRITILPQOYLRPVEDVATSDDCYKFAISQSTGTGAVIMEGFYV 405
DB 346 LYLMGEVTNOSRITILPQOYLRPVEDVATSDDCYKFAISQSTGTGAVIMEGFYV 405
QY 361 FDRARRIGFAVSAHVHDEFRTAAVEGPEVTLDMEDCGYNIPQDESLMTIATVMAAI 420
DB 361 FDRARRIGFAVSAHVHDEFRTAAVEGPEVTLDMEDCGYNIPQDESLMTIATVMAAI 420
QY 406 fdrarrigfavsachvhdefrtaaavgpftvclmedcgylnlptdeslmtiatvmaai 465
DB 406 fdrarrigfavsachvhdefrtaaavgpftvclmedcgylnlptdeslmtiatvmaai 465
QY 421 CALFMTPLCLMVCQWRCRLRQOHDFADDSLK 456
DB 421 CALFMTPLCLMVCQWRCRLRQOHDFADDSLK 456
QY 466 calfmplclmwcqwrclrlrqohdfadddlslk 501
DB 466 calfmplclmwcqwrclrlrqohdfadddlslk 501

RESULT 15
AA94769
ID AA94769 standard; Protein; 501 AA.
XX
AC AA94769;
XX
DT 12-FEB-2001 (first entry)
XX
DE Rat beta-secretase protein.
XX
KM Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease; rat;
KM Down's syndrome; amyloid angiodopathy; gene therapy; neuroprotective; ss.
OS Rattus sp.
XX
PN WO200058479-A1.
XX
PD 05-OCT-2000.
XX
PF 23-MAR-2000; 2000WO-US07755.

XX
PR 26-MAR-1999; 99US-0277229.
XX
PA (AMGE-) AMGEN INC.
PI Clifton M, Vassar RJ, Bennett BD;
XX
XX WPI: 2000-594643/56.
DR N-PSDB: AAA28280.
XX
PT Isolated beta-secretase nucleic acids and encoded polypeptides, useful
PT for diagnosis and gene therapy of Alzheimer's disease -
PS Claim 1; Fig 6; 145pp; English.
XX

This invention relates to 3 nucleotide sequences encoding beta-secretase
CC proteins. Beta-secretase is an enzyme involved in the production of one
CC of the components of amyloid plaques involved in Alzheimer's disease. The
CC invention includes an expression vector comprising the nucleotide
CC sequence, a host cell comprising the expression vector, and a process for
CC producing the protein through culturing the transformed cells. Also
CC included in the invention are a polypeptide derivative of the
CC beta-secretase protein, a fusion protein comprising beta-secretase fused
CC to a heterologous amino acid sequence, and a method for modulating the
CC levels of beta-secretase polypeptide in a mammal comprising administering
CC the polynucleotide sequence. Beta-secretase exhibits neuroprotective and
CC neurotropic activity. The beta-secretase gene and related genes on chromosomes
CC and as hybridization probes in diagnostic assays to test for the presence
CC of beta-secretase DNA or RNA, such as in Alzheimer's disease, Down's
CC syndrome, and amyloid angiodopathy. The nucleotide sequence may also be
CC used as anti-sense inhibitors of beta-secretase expression, in gene
CC therapy of Alzheimer's disease, and for the identification of compounds
CC that modulate beta-secretase activity. Antibodies to the beta-secretase
CC protein may be used for in vitro and in vivo diagnostic purposes to
CC detect the presence of beta-secretase polypeptide in a body fluid or cell
CC sample. The present sequence represents the rat beta-secretase protein.
SQ Sequence 501 AA;

Query Match 98.8%; Score 2392; DB 21; Length 501;
Best Local Similarity 98.8%; Pred. No. 2,9e-239;
Matches 448; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 ETDEPEEPGRGSGFVEMVDNLKSGSGGYVEMTVGSPPTQNLIVDTGSSNFAVGAAP 60
DB 46 etdepeepgrgsgfvevmdnlrgksqgyvemtvgspptqnlivdtgssnfavgaap 105
QY 61 HPFLHRYRQRLSTYRDLRKGYVVPYTOGKWESELTDLVSIHPGPNVTVRANIAITE 120
DB 106 hpfllhryrqrlstydrlrkgyvvpptogkwezeltdlvsihpgpnvtvraniaite 165
QY 121 SDRFFINGSMWEGILGLAVAEIARPDLSLEPFDSLVKOTHPVNLPSLQCGAGFPLNOS 180
DB 121 SDRFFINGSMWEGILGLAVAEIARPDLSLEPFDSLVKOTHPVNLPSLQCGAGFPLNOS 180
QY 166 Sdkffingsmwegilglavaelarppdslepfdsivkqthpnlsfqlcgagfplnds 225
DB 166 Sdkffingsmwegilglavaelarppdslepfdsivkqthpnlsfqlcgagfplnds 225
QY 181 EVLASVGSMTIGIDHSILYTGSLMTPTPIRREYVYVIVRVEINQDLMCKEYNDK 240
DB 181 EVLASVGSMTIGIDHSILYTGSLMTPTPIRREYVYVIVRVEINQDLMCKEYNDK 240
QY 226 EVLASVGSMTIGIDHSILYTGSLMTPTPIRREYVYVIVRVEINQDLMCKEYNDK 285
DB 226 EVLASVGSMTIGIDHSILYTGSLMTPTPIRREYVYVIVRVEINQDLMCKEYNDK 285
QY 241 SIVDSGTTNLRPKKVEFAVKSIRKASTKEKPPDFGFWGEOIYVQOACTPTPNTPIVPS 300
DB 241 SIVDSGTTNLRPKKVEFAVKSIRKASTKEKPPDFGFWGEOIYVQOACTPTPNTPIVPS 300
QY 286 SIVDSGTTNLRPKKVEFAVKSIRKASTKEKPPDFGFWGEOIYVQOACTPTPNTPIVPS 345
DB 286 SIVDSGTTNLRPKKVEFAVKSIRKASTKEKPPDFGFWGEOIYVQOACTPTPNTPIVPS 345
QY 301 LYLMGEVTNOSRITILPQOYLRPVEDVATSDDCYKFAISQSTGTGAVIMEGFYV 360
DB 301 LYLMGEVTNOSRITILPQOYLRPVEDVATSDDCYKFAISQSTGTGAVIMEGFYV 360
QY 346 LYLMGEVTNOSRITILPQOYLRPVEDVATSDDCYKFAISQSTGTGAVIMEGFYV 405
DB 346 LYLMGEVTNOSRITILPQOYLRPVEDVATSDDCYKFAISQSTGTGAVIMEGFYV 405
QY 361 FDRARRIGFAVSAHVHDEFRTAAVEGPEVTLDMEDCGYNIPQDESLMTIATVMAAI 420
DB 361 FDRARRIGFAVSAHVHDEFRTAAVEGPEVTLDMEDCGYNIPQDESLMTIATVMAAI 420
QY 406 fdrarrigfavsachvhdefrtaaavgpftvclmedcgylnlptdeslmtiatvmaai 465
DB 406 fdrarrigfavsachvhdefrtaaavgpftvclmedcgylnlptdeslmtiatvmaai 465

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:23:30 ; Search time 0.766585 Seconds

(without alignments)
501.389 Million cell updates/sec

Title: US-09-724-571-103

Perfect score: 20

Sequence: 1 VKMD 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	33	2 S23094	beta-amyloid prote
2	20	100.0	52	1 RUME	rubredoxin - Megas
3	20	100.0	57	2 F60045	Alzheimer's diseas
4	20	100.0	57	2 F60045	Alzheimer's diseas
5	20	100.0	57	2 G60045	Alzheimer's diseas
6	20	100.0	57	2 D60045	Alzheimer's diseas
7	20	100.0	57	2 A60045	Alzheimer's diseas
8	20	100.0	57	2 B60045	Alzheimer's diseas
9	20	100.0	74	2 F96670	Alzheimer's diseas
10	20	100.0	75	2 A97799	hypothetical prote
11	20	100.0	82	2 P00438	hypothetical prote
12	20	100.0	84	2 T27174	Alzheimer's diseas
13	20	100.0	86	2 D98327	hypothetical prote
14	20	100.0	86	2 A12955	cell division topo
15	20	100.0	90	2 AD3625	cell division inhi
16	20	100.0	92	2 T12859	hypothetical prote
17	20	100.0	93	2 T73016	hypothetical prote
18	20	100.0	94	2 JN0660	heat shock protein
19	20	100.0	94	2 S32105	chaperonin groes -
20	20	100.0	103	2 AC1582	thioredoxin [impor
21	20	100.0	103	2 A11228	thioredoxin [impor
22	20	100.0	104	2 H95206	thioredoxin [impor
23	20	100.0	104	2 H98071	thioredoxin reduct
24	20	100.0	105	2 PH1526	gamma-aminobutylic
25	20	100.0	127	2 H90315	conserved hypothet
26	20	100.0	134	2 G69444	hypothetical prote
27	20	100.0	140	2 A69445	hypothetical prote
28	20	100.0	142	2 E84059	hypothetical prote
29	20	100.0	145	2 F75189	hypothetical prote

30	20	100.0	146	2 C69136	hypothetical prote
31	20	100.0	147	2 T19162	hypothetical prote
32	20	100.0	149	2 T51471	farnesylated prote
33	20	100.0	152	2 T04479	cinnamyl alcohol d
34	20	100.0	152	2 H96734	probable isoprenyl
35	20	100.0	152	2 T48054	hypothetical prote
36	20	100.0	153	2 B84697	probable small hea
37	20	100.0	153	2 PNO564	von Willebrand fac
38	20	100.0	153	2 D90594	hypothetical prote
39	20	100.0	153	2 T05686	hypothetical prote
40	20	100.0	154	2 D71455	farnesylated prote
41	20	100.0	155	2 G96622	hypothetical prote
42	20	100.0	155	2 F75040	probable Heat shoc
43	20	100.0	159	2 AC2533	hypothetical prote
44	20	100.0	168	2 S73644	hypothetical prote
45	20	100.0	173	2 G64318	hypothetical prote

ALIGNMENTS

RESULT 1

S23094

beta-amyloid protein precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 03-May-1996

C:Accession: S23094

R:Kojima, S.; Omori, M., 1992

FEBS Lett. 304, 57-60, 1992

A>Title: Two-way cleavage of beta-amyloid protein precursor by multicatalytic protein

A:Reference number: S23094; MUID:92316198

A:Accession: S23094

A:Molecule type: protein

A:Residues: 1-33 <K03>

C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinas

Query Match

100.0%; Score 20; DB 2; Length 33;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
||||

DB 3 VKMD 6

RESULT 2

RUME

rubredoxin - Megaspheara elsdenti

C:Species: Megaspheara elsdenti

C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 24-Oct-1997

C:Accession: A00277

R:Baehmayer, H.; Yasunobu, K.T.; Peel, J.L.; Mayhew, S.

J. Biol. Chem. 243, 1022-1030, 1968

A>Title: Non-heme iron proteins. V. The amino acid sequence of rubredoxin from Peptos

A:Reference number: A00277; MUID:68161650

A:Accession: A00277

A:Molecule type: protein

A:Residues: 1-52 <BAC>

C:Superfamily: rubredoxin; rubredoxin homology

C:Keywords: electron transfer; iron; metalloprotein

F:3-48/Domain: rubredoxin homology <RUB>

F:6,9,38,41/Binding site: iron (Cys) #status predicted

Query Match

100.0%; Score 20; DB 1; Length 52;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
||||

DB 49 VKMD 52

RESULT 3

E60045
Alzheimer's disease amyloid beta/A4 protein precursor - sheep (fragment)
C:Species: Ovis sp. (sheep)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 28-Jul-1995
C:Accession: E60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog,
A:Reference number: A60045; MUID:92017079
A:Accession: E60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56130
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase 1
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 57;
100.0%; Pred. No. 54;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
|||||
DB 3 VKMD 6

RESULT 4
E60045
Alzheimer's disease amyloid beta/A4 protein precursor - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 13-Aug-1999
C:Accession: F60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog,
A:Reference number: A60045; MUID:92017079
A:Accession: F60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56127; NID:g1895; PIDN:CAA39592.1; PID:g1896
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase 1
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 57;
100.0%; Pred. No. 54;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
|||||
DB 3 VKMD 6

RESULT 5
G60045
Alzheimer's disease amyloid beta/A4 protein precursor - guinea pig (fragment)
C:Species: Cavia porcellus (guinea pig)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 28-Jul-1995
C:Accession: G60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog,
A:Reference number: A60045; MUID:92017079
A:Accession: G60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56126
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase 1
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 57;
100.0%; Pred. No. 54;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4

DB 3 VKMD 6
|||||

RESULT 6
D60045
Alzheimer's disease amyloid beta/A4 protein precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 28-Jul-1995
C:Accession: D60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d
A:Reference number: A60045; MUID:92017079
A:Accession: D60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56124
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 57;
100.0%; Pred. No. 54;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
|||||
DB 3 VKMD 6

RESULT 7
A60045
Alzheimer's disease amyloid beta/A4 protein precursor - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 28-Jul-1995
C:Accession: A60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d
A:Reference number: A60045; MUID:92017079
A:Accession: A60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56125
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 57;
100.0%; Pred. No. 54;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
|||||
DB 3 VKMD 6

RESULT 8
B60045
Alzheimer's disease amyloid beta/A4 protein precursor - polar bear (fragment)
C:Species: Ursus maritimus (polar bear)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 13-Aug-1999
C:Accession: B60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d
A:Reference number: A60045; MUID:92017079
A:Accession: B60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56128; NID:g2165; PIDN:CAA39593.1; PID:g2166
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 57;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
1111
Db 3 VKMD 6

RESULT 9

hypothetical protein F13011.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96670
R:Thillogis, A.; Ecker, J.R.; Palm, C.J.; Federpfeil, N.A.; Kaul, S.; White, O.; Alonso,
Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hutzlar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: F96670
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-74 <STO>
A:Cross-references: GB:AE005173; NID:95042411; PIDN:AD38250.1; GSPDB:GN00141
C:Genetics:
A:Gene: F13011.6
A:Map position: 1

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 74;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
1111
Db 13 VKMD 16

RESULT 10

A97799
hypothetical protein RC0793 [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: A97799
R: Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: A97799
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75 <KUR>
A:Cross-references: GB:AE006914; PIDN:AL03331.1; PID:G15619891; GSPDB:GN00173
C:Genetics:
A:Gene: RC0793

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
1111
Db 12 VKMD 15

RESULT 11

P00438

Alzheimer's disease amyloid A4 protein precursor - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C:Accession: P00438; C60045
R: Davidson, J.S.; West, R.L.; Kotikalapudi, P.; Maroun, L.E.
Biochem. Biophys. Res. Commun. 188, 905-911, 1992
A:Title: Sequence and methylation in the beta/A4 region of the rabbit amyloid precurs
A:Reference number: P00438; MUID:93075180
A:Accession: P00438
A:Molecule type: DNA
A:Residues: 1-82 <DAV>
A:Cross-references: GB:M83558; GB:M83657
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d
A:Reference number: A60045; MUID:92017079
A:Accession: C60045
A:Molecule type: mRNA
A:Residues: 12-68 <JOH>
A:Cross-references: EMBL:X56129
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type protein
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; Down's syndrome

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 82;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
1111
Db 14 VKMD 17

RESULT 12

T27174
hypothetical protein Y54G11A.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27174
R:Wallis, J.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z20322
A:Accession: T27174
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-84 <WIL>
A:Cross-references: EMBL:AL034488; NID:el359895; PIDN:CAA22454.1; CESP:Y54G11A.11
A:Experimental source: clone Y54G11A
C:Genetics:
A:Gene: CESP:Y54G11A.11
A:Introns: 39/2

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 84;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
1111
Db 37 VKMD 40

RESULT 13

D98327
cell division topological specificity factor [imported] - Agrobacterium tumefaciens (C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: D98327
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: D98327
A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-86 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK90142.1; PID:q15160139; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L_3134
 A:Map position: linear chromosome

Query Match 100.0%; Score 20; DB 2; Length 86;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
 ||||
 Db 60 VKMD 63

RESULT 14

AI2955
 cell division topological specificity factor mine [imported] - Agrobacterium tumefaciens

C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AI2955

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
 : Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AI2955

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-86 <KUR>

A:Cross-references: GB:AE008689; PIDN:AAL44063.1; PID:q17741627; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: mine

A:Map position: linear chromosome

Query Match 100.0%; Score 20; DB 2; Length 86;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
 ||||
 Db 60 VKMD 63

RESULT 15

AD3625
 cell division inhibitor mine [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AD3625

R:Delvecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 .; Marur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AD3625

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <KUR>

A:Cross-references: GB:AE008918; PIDN:AAL54167.1; PID:q17985133; GSPDB:GN00191

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI10925

A:Map position: II

Query Match 100.0%; Score 20; DB 2; Length 90;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
 ||||
 Db 63 VKMD 66

Search completed: October 30, 2002, 12:31:38
 Job time : 1.76658 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:10 ; Search time 0.34398 Seconds

(Without alignments)
450.253 Million cell updates/sec

Title: US-09-724-571-103

Perfect score: 20

Sequence: 1 VKMD 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	52	1 RUBR_MEGEL	P00271 megasphaera
2	20	100.0	57	1 A4_PIG	Q28023 sus scrofa
3	20	100.0	57	1 A4_URSM	Q28149 ursus marit
4	20	100.0	58	1 A4_CANFA	Q28280 canis famli
5	20	100.0	58	1 A4_RABIT	Q28748 oryctolagus
6	20	100.0	58	1 A4_SHEEP	Q28757 ovis aries
7	20	100.0	59	1 A4_BOVIN	Q28053 bos taurus
8	20	100.0	94	1 CH10_LACLA	P37283 lactococcus
9	20	100.0	103	1 TH10_LISMO	Q28386 listeria mo
10	20	100.0	134	1 VF60_ARCFU	Q28712 archaeoglob
11	20	100.0	140	1 VF62_ARCFU	Q28710 archaeoglob
12	20	100.0	168	1 VF24_MYCPN	P75254 mycoplasma
13	20	100.0	173	1 VF150_METJA	Q57614 methanococc
14	20	100.0	177	1 VP22_ASEB7	P23169 african swi
15	20	100.0	179	1 VAF4_CAEEL	P52882 caenorhabdi
16	20	100.0	194	1 RS7_FUGRU	P50894 fugu rubrip
17	20	100.0	196	1 VF01_MYCPN	P75286 mycoplasma
18	20	100.0	199	1 PHB1_HUMAN	P31941 homo sapien
19	20	100.0	209	1 HOXY_ALCEU	P23319 alcalligenes
20	20	100.0	224	1 COX2_ALSCO	P48889 albinaria c
21	20	100.0	224	1 COX2_ALBTU	Q09334 albinaria t
22	20	100.0	226	1 GSHP_BOVIN	P37141 bos taurus
23	20	100.0	226	1 GSHP_HUMAN	P23352 homo sapien
24	20	100.0	226	1 GSHP_MOUSE	P46412 mus musculu
25	20	100.0	226	1 GSHP_MOUSE	P23352 mus musculu
26	20	100.0	229	1 GSHP_MOUSE	P23352 mus musculu
27	20	100.0	230	1 COX2_PARLI	P12701 paracentrot
28	20	100.0	230	1 COX2_CARLU	Q78682 carassius a
29	20	100.0	230	1 COX2_CYPCA	P34189 cyprinus ca
30	20	100.0	230	1 COX2_CYPCA	P24987 cyprinus ca
31	20	100.0	230	1 COX2_ONCMY	P48171 oncorhynchu
32	20	100.0	230	1 COX2_SALSA	Q37677 salmo salar
33	20	100.0	230	1 COX2_SCYCA	Q79404 scyllorhinu
			230	1 COX2_SQUAC	Q92251 squallus aca

34	20	100.0	230	1 VNS1_IACKG	P30909 influenza a
35	20	100.0	230	1 VNS1_IADA2	P03501 influenza a
36	20	100.0	230	1 VNS1_IAMA6	P13137 influenza a
37	20	100.0	230	1 VNS1_IAP11	P13141 influenza a
38	20	100.0	230	1 VNS1_IAP13	P13143 influenza a
39	20	100.0	230	1 VNS1_IATK3	P30910 influenza a
40	20	100.0	230	1 VNS1_IATK3	P30911 influenza a
41	20	100.0	230	1 VNS1_IATK3	P30912 influenza a
42	20	100.0	231	1 VNS1_IATK3	P30913 influenza a
43	20	100.0	231	1 COX2_LATCH	O03848 latimeria c
44	20	100.0	239	1 VAI8_YEAST	P39548 saccharomyc
45	20	100.0	243	1 TONB_ENTAE	P58429 agrobacteri
					P46383 enterobacte

ALIGNMENTS

RESULT 1
RUBR_MEGEL STANDARD; PRT; 52 AA.
AC P00271.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Rubredoxin (rd).
OS Megasphaera elsdenii.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Sporomusa subbranch;
OC Megasphaera.
OX NCBI_TaxID=907;
RN [1]
RP SEQUENCE.
RX MEDLINE=68161650; PubMed=5640967;
RA Bachmayer H., Yasunobu K.T., Peel J.L., Mayhew S.G.;
RT "Non-heme iron proteins. V. The amino acid sequence of rubredoxin
RT from peptostreptococcus elsdenii.";
RL J. Biol. Chem. 243:1022-1030(1968).
-1- FUNCTION: RUBREDOXIN IS A SMALL, NONHEME, IRON PROTEIN LACKING
ACID-LABILE SULFIDE. ITS SINGLE FE, CHELATED TO 4CYS, FUNCTIONS
AS AN ELECTRON ACCEPTOR AND MAY ALSO STABILIZE THE CONFORMATION
OF THE MOLECULE.
CC -1- SIMILARITY: BELONGS TO THE RUBREDOXIN FAMILY.
CC PIR: A00277; RUME.
CC DR HSSP; P00269; IRB9.
CC DR InterPro; IPR004039; Rubredox.
CC DR InterPro; IPR001052; Rubredoxin.
CC DR Pfam; PF00301; rubredoxin.1.
CC DR PRINTS; PR00163; RUBREDOXIN.
CC DR ProDom; PD001610; Rubredoxn_domain; 1.
CC DR PROSITE; PS00202; RUBREDOXIN.1.
KW Electron transport; Iron; Metal-binding.
FT METAL 6 6
FT METAL 9 9
FT METAL 38 38
FT METAL 41 41
FT METAL 41 41
SQ SEQUENCE 52 AA; 5616 MW; A13AB077992PDPD4 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. NO. 30;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
DB 49 VKMD 52

RESULT 2
A4_PIG STANDARD; PRT; 57 AA.
AC Q29023;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid

```

DE protein (Beta-APP) (A-beta)] (Fragment).
GN APP.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis.";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -1- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: X56127; CAA39592.1; -.
CC DR HSSP: P05067; IBA4.
CC DR InterPro: IPR001868; A4_APP.
CC DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
CC DR PROSITE: PS00320; A4_INTRA; PARTIAL.
CC KW Glycoprotein; Amyloid; Neurone; Transmembrane.
CC FT NON_TER 1 1
CC FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
CC FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 34 57 POTENTIAL.
CC FT NON_TER 57 57
CC SQ SEQUENCE 57 AA; 6172 MW; 84209D88EBA82DFA CRC64;

Query Match 100.0%; Score 20; DB 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
   ||||
Db 3 VKMD 6

RESULT 3
A4_URSUMA STANDARD; PRT; 57 AA.
AC 029149;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [contains: Beta-amyloid
DE protein (Beta-APP) (A-beta)] (Fragment).
GN APP.
OS Ursus maritimus (Polar bear) (Thalarcos maritimus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Ursus.
OX NCBI_TaxID=29073;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis.";

```

```

RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -1- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL: X56128; CAA39593.1; -.
CC DR HSSP: P05067; IABL.
CC DR InterPro: IPR001868; A4_APP.
CC DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
CC DR PROSITE: PS00320; A4_INTRA; PARTIAL.
CC KW Glycoprotein; Amyloid; Neurone; Transmembrane.
CC FT NON_TER 1 1
CC FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
CC FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 34 57 POTENTIAL.
CC FT NON_TER 57 57
CC SQ SEQUENCE 57 AA; 6172 MW; 84209D88EBA82DFA CRC64;

Query Match 100.0%; Score 20; DB 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
   ||||
Db 3 VKMD 6

RESULT 4
A4_CANFA STANDARD; PRT; 58 AA.
AC 028280;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [contains: Beta-amyloid
DE protein (Beta-APP) (A-beta)] (Fragment).
GN APP.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis.";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -1- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----

```

```

CC -----
DR EMBL: X56125; CAA39590.1; -
DR HSSP: P05067; 1BA4.
DR InterPro: IPR001868; A4_APP.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
DR Glycoprotein: Amyloid; Neurone; Transmembrane.
KW NON_TER 1 1
FT CHAIN 7 49 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 34 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 35 58 POTENTIAL.
FT NON_TER 58 58
SQ SEQUENCE 58 AA: 6285 MW: 8469D488A2E12DFA CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 58;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
   ||||
Db 4 VKMD 7

RESULT 5
A4_RABIT STANDARD: PRT; 58 AA.
AC Q28748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
DE protein (beta-Ap) (A-beta)] (Fragment).
GN App.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis.";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -1- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X56129; CAA39594.1; -
DR HSSP: P05067; 1BA4.
DR InterPro: IPR001868; A4_APP.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
KW Glycoprotein: Amyloid; Neurone; Transmembrane.
KW NON_TER 1 1
FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 34 57 POTENTIAL.
FT DOMAIN 58 >58 CYTOPLASMIC (POTENTIAL).
FT NON_TER 58 58
SQ SEQUENCE 58 AA: 6300 MW: F434209D88BBA82D CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 58;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
   ||||
Db 3 VKMD 6

RESULT 6
A4_SHEEP STANDARD: PRT; 58 AA.
AC Q28757;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
DE protein (beta-Ap) (A-beta)] (Fragment).
GN App.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_Taxid=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis.";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -1- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X56130; CAA39595.1; -
DR HSSP: P05067; 1AHL.
DR InterPro: IPR001868; A4_APP.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
KW Glycoprotein: Amyloid; Neurone; Transmembrane.
KW NON_TER 1 1
FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 34 57 POTENTIAL.
FT DOMAIN 58 >58 CYTOPLASMIC (POTENTIAL).
FT NON_TER 58 58
SQ SEQUENCE 58 AA: 6300 MW: F434209D88BBA82D CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 1; Length 58;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
   ||||
Db 3 VKMD 6

RESULT 7
A4_BOVIN

```

```

ID      A4-BOVIN          STANDARD:          PRT:          59 AA.
AC      Q28053;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Alzheimer's disease amyloid A4 protein homolog [contains: Beta-amyloid
DE      protein (Beta-APP) (A-beta)] (Fragment).
GN      APP.
OS      Bos taurus (Bovine).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC      Bovidae; Bovinae; Bos.
OX      NCBI_TaxID=9913;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RX      MEDLINE=92017079; PubMed=1656157;
RT      Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT      "Conservation of the sequence of the Alzheimer's disease amyloid
RT      peptide in dog, polar bear and five other mammals by cross-species
RT      polymerase chain reaction analysis.";
RL      Brain Res. Mol. Brain Res. 10:299-305(1991).
CC      -1- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC      INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC      G1O (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- SIMILARITY: BELONGS TO THE APP FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X56124; CA39589.1; -
DR      EMBL: X56126; CA39591.1; -
DR      HSSP: P05067; 1BA4.
DR      InterPro: IPR001868; A4_APP.
DR      PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR      PROSITE: PS00320; A4_INTRA; PARTIAL.
KW      Glycoprotein; Amyloid; Neurone; Transmembrane.
FT      NON_TER          1          1
FT      CHAIN            1          49      BETA-AMYLOID PROTEIN (POTENTIAL).
FT      DOMAIN          <1          34      EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM        35          58      POTENTIAL.
FT      DOMAIN          59          59      CYTOPLASMIC (POTENTIAL).
FT      NON_TER          59          59
SQ      SEQUENCE      59 AA; 6414 MW; F43469DA8A2E12D CRC64;

Query Match      100.0%; Score 20; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 34;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

OY      1 VKMD 4
        ||||
DB      4 VKMD 7

RESULT 8
CH10_LACLA          STANDARD:          PRT:          94 AA.
ID      CH10_LACLA
AC      P37283;
DT      01-OCT-1994 (Rel. 30, Created)
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      10 kDa chaperonin (Protein Cpn10) (groES protein).
GN      GROES OR GROES OR IL0393.
OS      Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC      Lactococcus.
OX      NCBI_TaxID=1360;

```

```

RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93252268; PubMed=8486277;
RA      Kim S.G., Batt C.A.;
RT      "Cloning and sequencing of the Lactococcus lactis subsp. lactis
RT      groESL operon.";
RL      Gene 127:121-126(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=IL1403;
RX      MEDLINE=21235186; PubMed=11337471;
RA      Bolotin A., Wincker P., Manger S., Jallou O., Malarne K.,
RA      Weissenbach J., Ehrlich S.D., Sorokin A.;
RT      "The complete genome sequence of the lactic acid bacterium Lactococcus
RT      lactis ssp. lactis IL1403.";
RL      Genome Res. 11:731-753(2001).
CC      -1- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
CC      THE ATPASE ACTIVITY OF THE LATTER.
CC      -1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
CC      (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X71132; CA50445.1; -
DR      EMBL: AE006276; AAK0491.1; -
DR      PIR: S32105; S32105.
DR      PIR: JN0660; JN0660.
DR      InterPro: IPR001476; Cpn10.
DR      Pfam: PF00166; Cpn10; 1.
DR      PRINTS: PR00297; CHAPERONIN10.
DR      ProDom: PD00566; Cpn10; 1.
DR      PROSITE: PS00681; CHAPERONINS_CPN10; 1.
KW      Chaperone; Complete proteome.
SQ      SEQUENCE      94 AA; 10221 MW; 291913B147CA66B CRC64;

Query Match      100.0%; Score 20; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 54;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

OY      1 VKMD 4
        ||||
DB      74 VKMD 77

RESULT 9
TH10_LISMO          STANDARD:          PRT:          103 AA.
ID      TH10_LISMO
AC      Q95386;
DT      01-MAR-2002 (Rel. 41, Created)
DT      01-MAR-2002 (Rel. 41, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      Thioedoxin (Trx).
GN      TRXA OR LMO1233 OR LIN1196.
OS      Listeria monocytogenes, and
OS      Listeria innocua.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group;
OC      Bacillus/Staphylococcus group; Listeria.
OX      NCBI_TaxID=1639, 1642;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      SPECIES=L.monocytogenes; STRAIN=BGD;
RA      Tarchanov M., Borovok I., Aharonowitz Y., Cohen G.;
RT      "Isolation, cloning and characterization of the Listeria monocytogenes
RT      thioedoxin gene, trxa.";
RL      Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.
RN      [2]

```

```

RP SEQUENCE FROM N.A.
RC SPECIES=L. monocytogenes, and L. innocua;
RX STRAIN=ED-e / Serovar 1/2a, and CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed-11679669;
RA Glaeser P., Frangoul L., Buchrieser C., Rusnlok C., Amend A.,
RA Baquero F., Berche P., Bioecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Eutlier K.-D., Fsihl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hahn T., Hauf J., Jackson D.,
RA Jones L.-M., Kereit U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madano E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstedt G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tietze A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RX "Comparative genomics of Listeria species.";
RX Science 294:849-852(2001).
CC -1- FUNCTION: COMPONENT OF THE THIOREDOXIN-THIOREDOXIN REDUCTASE
CC SYSTEM. THIOREDOXIN PARTICIPATES IN VARIOUS REDOX REACTIONS
CC THROUGH THE REVERSIBLE OXIDATION OF ITS ACTIVE CENTER DITHIOL, TO
CC A DISULFIDE, AND CATALYZES DITHIOL-DISULFIDE EXCHANGE REACTIONS
CC (By similarity).
CC -1- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ13306; CAB40815.2; -
DR EMBL: AL591978; CAC9311.1; -
DR EMBL: AL596167; CAC96427.1; -
DR HSSP: P80579; 10WU.
DR Listlist: LIND1196; -
DR Listlist: LIND1233; -
DR InterPro: IPR000063; Thired.
DR Pfam: PF00085; Thired; 1.
DR PRINTS: PR00421; THIOREDOXIN.
DR PROSITE: PS00194; THIOREDOXIN; 1.
DR Oxidoreductase: Redox-active center: Complete proteome.
KW DISULFID 28 31 REDOX-ACTIVE (BY SIMILARITY).
FT SEQUENCE 103 AA: 11620 MW; 01F6A77434559A6 CRC64;
SQ
Query Match 100.0%; Score 20; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VKMD 4
DB 52 VKMD 55

```

```

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervlavage A.R., Graham D.E., Kyriakides N.C.,
RA Fleischman R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.R., Badger J.H., Glodok A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RX "The complete genome sequence of the hyperthermophilic, sulphate-
RX reducing archaeon Archaeoglobus fulgidus.";
RX Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000995; AAB89698.1; -
DR TIGR: AF1560; -
KW Hypothetical protein: Transmembrane: Complete proteome.
FT TRANSMEM 13 35 POTENTIAL.
SQ SEQUENCE 134 AA: 15036 MW; EEB83D22A904A79 CRC64;
Query Match 100.0%; Score 20; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VKMD 4
DB 83 VKMD 86

```

```

RESULT 10
YF62_ARCFU STANDARD: PRT; 134 AA.
ID YF62_ARCFU STANDARD: PRT; 134 AA.
AC 028710;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF1560.
GN AF1560.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RX Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

```

```

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervlavage A.R., Graham D.E., Kyriakides N.C.,
RA Fleischman R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.R., Badger J.H., Glodok A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RX "The complete genome sequence of the hyperthermophilic, sulphate-
RX reducing archaeon Archaeoglobus fulgidus.";
RX Nature 390:364-370(1997).
CC -----
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000994; AAB89687.1; -
CC DR TIGR; AF1562; -
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 20 42 POTENTIAL.
CC FT TRANSMEM 88 110 POTENTIAL.
CC FT TRANSMEM 115 137 POTENTIAL.
CC FT TRANSMEM 140 AA; 15667 MW; 937DCB5585A17991 CRC64;
CC SEQUENCE
CC
CC Query Match 100.0%; Score 20; DB 1; Length 140;
CC Best Local Similarity 100.0%; Pred. No. 80;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC OY 1 VKMD 4
CC DB 2 VKMD 5
CC
CC RESULT 12
CC YF24_MYCPN STANDARD; PRT; 168 AA.
CC ID YF24_MYCPN
CC AC P75254;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Hypothetical protein MPN524 (G12,orf168).
CC GN MPN524 OR MP318.
CC OS Mycoplasma pneumoniae.
CC OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
CC OC Mycoplasmataceae; Mycoplasma.
CC OX NCBI_TaxID=2104;
CC [1]
CC RP SEQUENCE FROM N.A.
CC RX STRAIN-ATCC 29342 / M129;
CC RX MEDLINE=97105885; PubMed=8948633;
CC RA Himmelfeich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
CC RA Herrmann R.;
CC RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
CC pneumoniae."
CC RL Nucleic Acids Res. 24:4420-4449(1996).
CC CC -1- SIMILARITY: BELONGS TO THE UPF0134 FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC CC
CC DR EMBL; AE000029; AAB95966.1; -
CC DR InterPro; IPR002862; DUF16.
CC DR Pfam; PF01519; DUF16; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 168 AA; 20110 MW; 0878B81DEB019170 CRC64;
CC
CC Query Match 100.0%; Score 20; DB 1; Length 168;
CC Best Local Similarity 100.0%; Pred. No. 95;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC OY 1 VKMD 4
CC DB 125 VKMD 128
CC
CC RESULT 13
CC Y150_METJA STANDARD; PRT; 173 AA.
CC ID Y150_METJA
CC AC Q57614;

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0150.
GN MJ0150.
OS Methanococcus jannaschii.
CC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
CC Methanococcus.
CC NCBI_TaxID=2190;
CC [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
CC RX MEDLINE=96337999; PubMed=8688087;
CC RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
CC RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
CC RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
CC RA Scott J.L., Geoghegan N.S.M., Feldman J.F., Fuhrmann J.L., Nguyen D.,
CC RA Uitterlisch T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
CC RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
CC RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
CC RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
CC jannaschii."
CC RL Science 273:1058-1073(1996).
CC CC -1- SIMILARITY: BELONGS TO THE M_JANNASCHII MJ0150 / MJ0739 / MJ0745 /
CC MJ1460 / MJ1642 FAMILY AND TO THE C-TERMINAL OF M_JANNASCHII
CC MJ0855.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC CC
CC DR EMBL; U67472; AAB98141.1; -
CC DR TIGR; MJ0150; -
CC DR InterPro; IPR004096; VAR.
CC DR Pfam; PF02830; VAR; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 173 AA; 19524 MW; A273F382C91D0315 CRC64;
CC
CC Query Match 100.0%; Score 20; DB 1; Length 173;
CC Best Local Similarity 100.0%; Pred. No. 98;
CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC OY 1 VKMD 4
CC DB 4 VKMD 7
CC
CC RESULT 14
CC VP22_ASFB7 STANDARD; PRT; 177 AA.
CC ID VP22_ASFB7
CC AC P23169;
CC DT 01-NOV-1991 (Rel. 20, Created)
CC DT 01-NOV-1991 (Rel. 20, Last sequence update)
CC DT 01-AUG-1992 (Rel. 23, Last annotation update)
CC DE Protein P22 precursor (Protein K177).
CC OS African swine fever virus (strain BA71V) (ASFV).
CC OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae;
CC CC African swine fever-like viruses.
CC NCBI_TaxID=10498;
CC [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=90219205; PubMed=2325203;
CC RA Gonzalez A., Calvo V., Almazan F., Almendral J.M., Ramirez J.C.,
CC RA de la Vega I., Blasco R., Vinuela E.;
CC RT "Multigene families in African swine fever virus: family 360."
CC J. Virol. 64:2073-2081(1990).
CC [2]
CC RP CHARACTERIZATION.

```


Search completed: October 30, 2002, 12:27:56
Job time : 2.34398 secs

RX MEDLINE-91134988: PubMed-1994575;

RA Camacho A., Vinuela E.;

RT "Protein p22 of African swine fever virus: an early structural
protein that is incorporated into the membrane of infected cells.";

RL Virology 181:251-257(1991).

CC -I- FUNCTION: EARLY STRUCTURAL PROTEIN THAT IS INCORPORATED INTO THE
MEMBRANE OF INFECTED CELLS.

CC -I- SUBCELLULAR LOCATION: PROBABLE EXTERNAL VIRAL STRUCTURAL PROTEIN.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U18466; AAA65240.1; -;

DR EMBL: M57546; AAA42681.1; -;

KW Early protein; Signal; Membrane; Structural protein.

FT SIGNAL

FT CHAIN 1 27 POTENTIAL.

FT SEQUENCE 177 AA; 20160 MW; 7C2EE66F13E762D1 CRC64;

Query Match Best Local Similarity 100.0%; Score 20; DB 1; Length 177;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4

Db 65 VKMD 68

RESULT 15

YAF4_CAEEL

ID YAF4_CAEEL STANDARD; PRT; 179 AA.

AC P52882;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Hypothetical 20.4 kDa protein F46C5.4 in chromosome II.

GN F46C5.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Thomas K.;

RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: 254281; CA91046.1; -;

DR Wormpep: F46C5.4; CE03346.

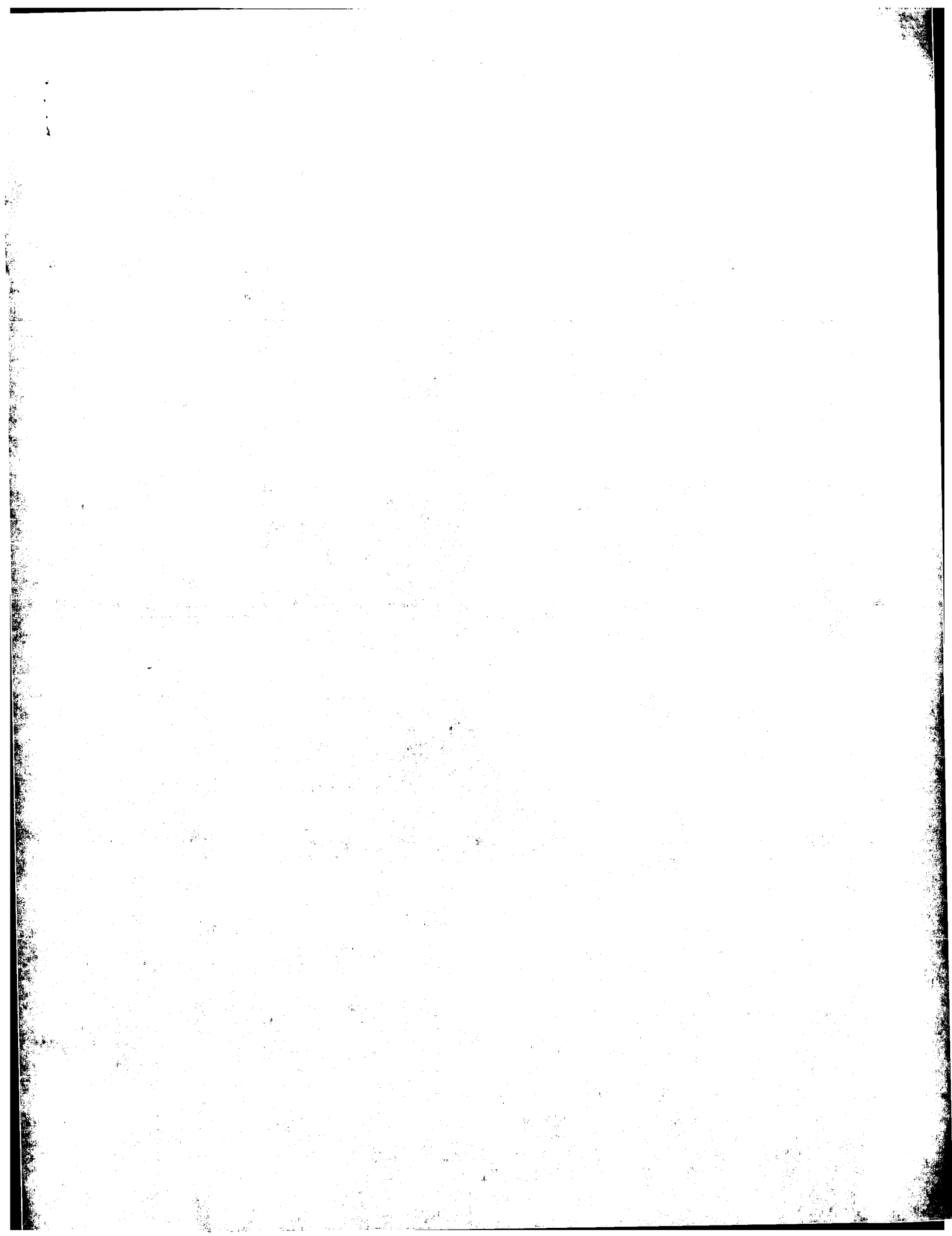
KW Hypothetical protein.

FT SEQUENCE 179 AA; 20375 MW; 4662982743926408 CRC64;

Query Match Best Local Similarity 100.0%; Score 20; DB 1; Length 179;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4

Db 119 VKMD 122



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      October 30, 2002, 12:21:30 ; Search time 1.24816 Seconds
              (without alignments)
              554.401 Million cell updates/sec
```

Title: US-09-724-571-103
Perfect score: 20

Sequence: 1 VKMD 4

Sequence: 1 VKMD 4

Scoring table: BLOSUM62

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```

1:  sp_archae_19:*
   sp_archae_19:sp archaeal
2:  sp_bacteria:*
   sp_bacteria:sp bacterial
3:  sp_fungi:*
   sp_fungi:sp fungal
4:  sp_human:*
   sp_human:sp human
5:  sp_invertebrate:*
   sp_invertebrate:sp invertebrate
6:  sp_mammal:*
   sp_mammal:sp mammal
7:  sp_minc:*
   sp_minc:sp minic
8:  sp_organelle:*
   sp_organelle:sp organelle
9:  sp_phage:*
   sp_phage:sp phage
10: sp_plant:*
    sp_plant:sp plant
11: sp_podent:*
    sp_podent:sp podent
12: sp_virus:*
    sp_virus:sp virus
13: sp_unclassified:*
    sp_unclassified:sp unclassified
14: sp_unclassified:*
    sp_unclassified:sp unclassified
15: sp_virus:*
    sp_virus:sp virus
16: sp_bacteriap:*
    sp_bacteriap:sp bacteriap
17: sp_archaeap:*
    sp_archaeap:sp archaeap

```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	60	5	Q9BMS3	Q9bms3 plasmodium
2	20	100.0	62	5	Q9BJU7	Q9bjj7 plasmodium
3	20	100.0	65	5	Q9BJU4	Q9bjj4 plasmodium
4	20	100.0	65	11	O546596	O546596 mus muscu
5	20	100.0	74	10	Q9XTR8	Q9xtr8 arabidops
6	20	100.0	75	16	Q92HH8	Q92hh8 rickettsi
7	20	100.0	75	11	O35463	O35463 cricetulu
8	20	100.0	82	4	P78438	P78438 homo sapie
9	20	100.0	82	4	Q9UI77	Q9ui77 homo sapie
10	20	100.0	82	4	O16014	O16014 homo sapie
11	20	100.0	82	4	O16019	O16019 homo sapie
12	20	100.0	82	4	O16030	O16030 homo sapie
13	20	100.0	82	5	O9U501	O9u501 homo sapie
14	20	100.0	83	4	O96IT4	O96it4 manduca se
15	20	100.0	84	5	O9XVZ8	O9xvz8 caenorhabd
16	20	100.0	86	1	Q9C4V4	Q9c4v4 sulliobus

17	20	100.0	90	11	P70443
18	20	100.0	9	0	064108
19	20	100.0	92	16	Q34730
20	20	100.0	93	5	Q21047
21	20	100.0	94	2	Q9AEP8
22	20	100.0	98	5	Q9P45
23	20	100.0	101	12	Q91ZR8
24	20	100.0	103	16	Q9S386
25	20	100.0	104	5	Q8GP75
26	20	100.0	104	5	Q8GP72
27	20	100.0	104	5	Q9GN84
28	20	100.0	104	5	Q9GN83
29	20	100.0	104	5	Q9GN82
30	20	100.0	104	16	Q97P68
31	20	100.0	105	5	Q9GP79
32	20	100.0	105	5	Q8GP78
33	20	100.0	105	5	Q9GN12
34	20	100.0	106	5	Q9GP80
35	20	100.0	115	15	Q74763
36	20	100.0	125	12	Q9J0E5
37	20	100.0	127	17	Q97XY5
38	20	100.0	137	12	Q83000
39	20	100.0	137	12	Q91E71
40	20	100.0	137	12	Q91E36
41	20	100.0	138	10	Q9ZRD7
42	20	100.0	142	12	Q91Z44
43	20	100.0	145	16	Q9K7K4
44	20	100.0	145	12	Q91MX8
45	20	100.0	145	17	Q9VZN3
					P70443 mus musculus
					Q64108 bacterioph
					Q34730 bacillus su
					Q21047 caenorhabdi
					Q9AEP8 lactococcus
					Q9P45 drosophila
					Q91ZR8 tick-borne
					Q9S386 listeria mo
					Q9GP75 drosophila
					Q9GP72 drosophila
					Q9GN84 drosophila
					Q9GN83 drosophila
					Q9GN82 drosophila
					Q97P68 streptococc
					Q9GP79 drosophila
					Q9GP78 drosophila
					Q9GN12 drosophila
					Q9GP80 drosophila
					Q74763 human immun
					Q9J0E5 tick-borne
					Q97XY5 sulfobus
					Q83000 bunyavirus
					Q91E71 tick-borne
					Q91E36 phocid herp
					Q9ZRD7 glycine max
					Q91Z44 tick-borne
					Q9K7K4 bacillus ha
					Q91MX8 lumby skin
					Q9VZN3 pyroccus

ALIGNMENTS

RESULT 1			
O9BMS3			
ID	O9BMS3	PRELIMINARY:	PRT: 60 AA.
AC	O9BMS3.		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)		
DE	ERYTHROCYTE MEMBRANE PROTEIN 1 (FRAGMENT).		
GN	VARL.		
OS	Plasmodium falciparum.		
CC	Eukaryota, Alveolata;		
OX	NCBI_taxid=583;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21273141; Pubmed=11378202;		
RA	Noviyanti R., Brown G.V., Wickham M.E., Duffly M.F., Cowman A.F.,		
RA	Reeder J.C.;		
RT	"Multiple var gene transcripts are expressed in Plasmodium falciparum		
RT	infected erythrocytes selected for adhesion.";		
RL	Mol. Biochem. Parasitol. 114:227-237(2001).		
DR	EMBL: AF306406; AAK00867.1; -		
FT	NON TER	1	
FT	NON TER	1	
SO	SEQUENCE	60 AA;	6798 MW; 8377668E9DB9291 CRC64;

Query Match	100.0%	Score 20;	DB 5;	Length 60;
Best Local Similarity	100.0%;	Pred. NO. 1.9e+02;		
Matches	4;	Conservative	0;	Mismatches 0;
			Indels	0;
			Caps	0

QY	1	VKMD	4
Db	6	VKMD	9

RESULT 2	
Q9BJJ6	
ID Q9BJJ6	PRELIMINARY;
AC Q9BJJ6;	PRT;
DT 01-JUN-2001 (TREMblrel, 17, Created)	62 AA

```

DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, last annotation update)
DE VARIANT SURFACE PROTEIN (FRAGMENT).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=93;
RA Nogueira P.A., Wunderlich G., Tada M.S., Costa J.D.N., Menezes M.J.,
  Scherf A., Pereira da Silva L.H.;
  "Plasmodium falciparum: repertoire of expressed var genes and adhesion
  properties to endothelial receptors of clinical isolates from patients
  in Rondonia (Brazilian western Amazon region).";
  Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF333282; AAK19553.1; -.
DR InterPro: IPR001230; Prenyltin.
DR PROSITE: PS00294; PRENYLTATION; UNKNOWN_1.
FT NON_TER 1 62
FT NON_TER 1 1
SQ SEQUENCE 62 AA; 6953 MW; 83209EA45C1D57A5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 5; Length 62;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
   ||||
DB 10 VKMD 13

RESULT 3
ID Q9BJJ4 PRELIMINARY; PRT; 62 AA.
AC Q9BJJ4;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, last annotation update)
DE VARIANT SURFACE PROTEIN (FRAGMENT).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=95;
RA Nogueira P.A., Wunderlich G., Tada M.S., Costa J.D.N., Menezes M.J.,
  Scherf A., Pereira da Silva L.H.;
  "Plasmodium falciparum: repertoire of expressed var genes and adhesion
  properties to endothelial receptors of clinical isolates from patients
  in Rondonia (Brazilian western Amazon region).";
  Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF333284; AAK19555.1; -.
DR InterPro: IPR001230; Prenyltin.
DR PROSITE: PS00294; PRENYLTATION; UNKNOWN_1.
FT NON_TER 1 62
FT NON_TER 1 1
SQ SEQUENCE 62 AA; 7029 MW; F55E3A8BAFADAC CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 5; Length 62;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
   ||||
DB 10 VKMD 13

RESULT 4
ID Q54696 PRELIMINARY; PRT; 65 AA.
AC Q54696;

```

```

DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE PROSTAGLANDIN F2 ALPHA RECEPTOR REGULATORY PROTEIN (FRAGMENT).
GN PGFRN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97109715; Pubmed=8951995;
RX Orlicky D.J., Nordeen S.K.;
  "Cloning, sequencing and proposed structure for a prostaglandin F2
  alpha receptor regulatory protein.";
  prostaglandins Leukot. Essent. Fatty Acids 55:261-268(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Orlicky D.J.;
  "Mouse prostaglandin F2 alpha receptor regulatory protein cDNA
  fragment.";
  Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF006201; AAB87643.1; -.
DR MGD: MGI:1277114; Pgfrn.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 65 AA; 7623 MW; 69D3E97775D86729 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 11; Length 65;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
   ||||
DB 7 VKMD 10

RESULT 5
ID Q9XIR8 PRELIMINARY; PRT; 74 AA.
AC Q9XIR8;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE F13011.6 PROTEIN.
DE F13011.6.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
  euosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
  Altati H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
  Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
  Lueros S., Schwartz J., Shilm P., Toriumi M., Vysotskaya V.S.,
  Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
  Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AC006193; AAD38250.1; -.
DR PROSITE: PS00294; PRENYLTATION; UNKNOWN_1.
SQ SEQUENCE 74 AA; 8859 MW; F76AAF00BE8818A CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 10; Length 74;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
   ||||
DB 13 VKMD 16

RESULT 6

```

```

092HH8
ID 092HH8 PRELIMINARY; PRT; 75 AA.
AC 092HH8;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DE 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HYPOTHEMETICAL PROTEIN RC0793.
GN RC0793.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiaseae; Rickettsia.
OX NCBI_Taxid=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALISH 7.
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
DR EMBL; AE008635; AL03331.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 75 AA; 8841 MW; 66EE9C73C0D29592 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 16; Length 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
DB 12 VKMD 15

RESULT 7
035463
ID 035463 PRELIMINARY; PRT; 79 AA.
AC 035463;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ALZHEIMER'S AMYLOID BETA PROTEIN (FRAGMENT).
GN BETA APP.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_Taxid=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Sambamurti K., Plunik I., Gandhi S.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF030413; AAB86608.1; -
DE HSSP; P05067; IBA4.
FT NON_TER 1 79
FT NON_TER 1 79
SQ SEQUENCE 79 AA; 8538 MW; 37F2C6C3BFF3F597 CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 11; Length 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
DB 18 VKMD 21

RESULT 8
P78438
ID P78438 PRELIMINARY; PRT; 82 AA.
AC P78438;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)

```

```

DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE AMYLOID PROTEIN (BETA-AMYLOID PROTEIN) (FRAGMENT).
GN APP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89392030; PubMed=2675837;
RA Johnstone E.M., Chaney M.O., Moore R.E., Ward K.E., Norris F.H.,
RA Little S.P.;
RT "Alzheimer's disease amyloid peptide is encoded by two exons and shows
RT similarity to soybean trypsin inhibitor.";
RL Biochem. Biophys. Res. Commun. 163:1248-1255(1989).
RN [2]
RP SEQUENCE OF 19-48 FROM N.A.
RX MEDLINE=87120329; PubMed=2949367;
RA Tanzi R.E., Gusella J.F., Watkins P.C., Bruns G.A., George-Hyslop P.,
RA Van Keuren M.L., Patterson D., Pagan S., Kurnit D.M., Neve R.L.;
RT "Amyloid beta protein gene: cDNA, mRNA distribution, and genetic
RT linkage near the Alzheimer locus.";
RL Science 235:880-884(1987).
RN [3]
RP SEQUENCE OF 32-63 FROM N.A.
RX MEDLINE=93035397; PubMed=1415269;
RA Kamino K., Orr H.T., Payaml H., Wijsman E.M., Alonso M.E., Pulsat S.M.,
RA Anderson L., O'dahl S., Nemens E., White J.A.;
RT "Linkage and mutational analysis of familial Alzheimer disease
RT kindreds for the APP gene region.";
RL Am. J. Hum. Genet. 51:998-1014(1992).
DR EMBL; M29270; AAA51768.1; -
DR EMBL; M29269; AAA51768.1; JOINED.
DR EMBL; M15532; AAA51564.1; -
DR EMBL; S45136; AAB23646.1; -
DR HSSP; P05067; IBA4.
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 82 AA; 8994 MW; 8DA9E42B813A070E CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 82;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VKMD 4
DB 14 VKMD 17

RESULT 9
090177
ID 090177 PRELIMINARY; PRT; 82 AA.
AC 090177;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE PRO0214 PROTEIN (DJ319D22.2).
GN DJ319D22.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Yu Y., Zhang C., Luo L., Ouyang S., Zhang S., Li W., Wu J., Zhou S.,
RA Liu M., He F.;
RT "Functional prediction of the coding sequences of 50 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Williams S.;
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.

```

DR EMBL: AF090903: AAF24027.1; -
 DR EMBL: AL133262: CAC08558.1; -
 SQ SEQUENCE 82 AA: 9443 MW: E29D37E1E2B327A1 CRC64;

Query Match 100.0%; Score 20; DB 4; Length 82;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
 ||||
 DB 19 VKMD 22

RESULT 10

Q16014 PRELIMINARY; PRT; 82 AA.
 AC Q16014;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BETA-AMYLOID PEPTIDE (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93236601; PubMed=8476439;
 RA Denman R.B., Rosenzweig R., Miller D.L.;
 RT "A system for studying the effect(s) of familial Alzheimer disease
 mutations on the processing of the beta-amyloid peptide precursor.";
 RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
 DR EMBL: S60721; AAB26263.2; -
 DR HSSP: P05067; 1BA4.
 FT NON_TER 1 82
 SQ SEQUENCE 82 AA: 8972 MW: F534AA5B3EA9230A CRC64;

Query Match 100.0%; Score 20; DB 4; Length 82;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
 ||||
 DB 15 VKMD 18

RESULT 11

Q16019 PRELIMINARY; PRT; 82 AA.
 ID Q16019;
 AC Q16019;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BETA-AMYLOID PEPTIDE (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93236601; PubMed=8476439;
 RA Denman R.B., Rosenzweig R., Miller D.L.;
 RT "A system for studying the effect(s) of familial Alzheimer disease
 mutations on the processing of the beta-amyloid peptide precursor.";
 RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
 DR EMBL: S61380; AAB26264.2; -
 DR HSSP: P05067; 1BA4.
 FT NON_TER 1 82
 SQ SEQUENCE 82 AA: 8938 MW: F534AA50E579230A CRC64;

Query Match 100.0%; Score 20; DB 4; Length 82;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
 ||||
 DB 15 VKMD 18

RESULT 12

Q16020 PRELIMINARY; PRT; 82 AA.
 ID Q16020;
 AC Q16020;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BETA-AMYLOID PEPTIDE (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93236601; PubMed=8476439;
 RA Denman R.B., Rosenzweig R., Miller D.L.;
 RT "A system for studying the effect(s) of familial Alzheimer disease
 mutations on the processing of the beta-amyloid peptide precursor.";
 RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
 DR EMBL: S61383; AAB26263.2; -
 DR HSSP: P05067; 1BA4.
 FT NON_TER 1 82
 SQ SEQUENCE 82 AA: 8882 MW: F534AA5AE5D9230A CRC64;

Query Match 100.0%; Score 20; DB 4; Length 82;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
 ||||
 DB 15 VKMD 18

RESULT 13

Q9U501 PRELIMINARY; PRT; 82 AA.
 ID Q9U501;
 AC Q9U501;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 9.5 KDA PROTEIN.
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Spingioidea; Sphingidae; Sphinginae; Manduca.
 NX NCBI_TaxID=7130;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20099029; PubMed=10620045;
 RA Robertson H.M., Martos R., Sears C., Todres E.Z., Walden K.O.,
 RA Nard J.B.;
 RT "Diversity of odourant binding proteins revealed by an expressed
 sequence tag project on male Manduca sexta moth antennae.";
 RL Insect Mol. Biol. 8:501-518(1999).
 DR EMBL: AF117587; AAF16709.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 82 AA: 9460 MW: AB889CE023AB4E5C CRC64;

Query Match 100.0%; Score 20; DB 5; Length 82;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
||||
DB 37 VKMD 40

RESULT 14

Q96IT4 PRELIMINARY; PRT; 83 AA.
AC Q96IT4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SIMILAR TO CG6244 GENE PRODUCT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE; AND RHABDOMYOSARCOMA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007516; AA07516.1; -
SO SEQUENCE 83 AA; 9462 MW; 77F1EFBE72AC6F CRC64;

Query Match 100.0%; Score 20; DB 4; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
||||
DB 37 VKMD 40

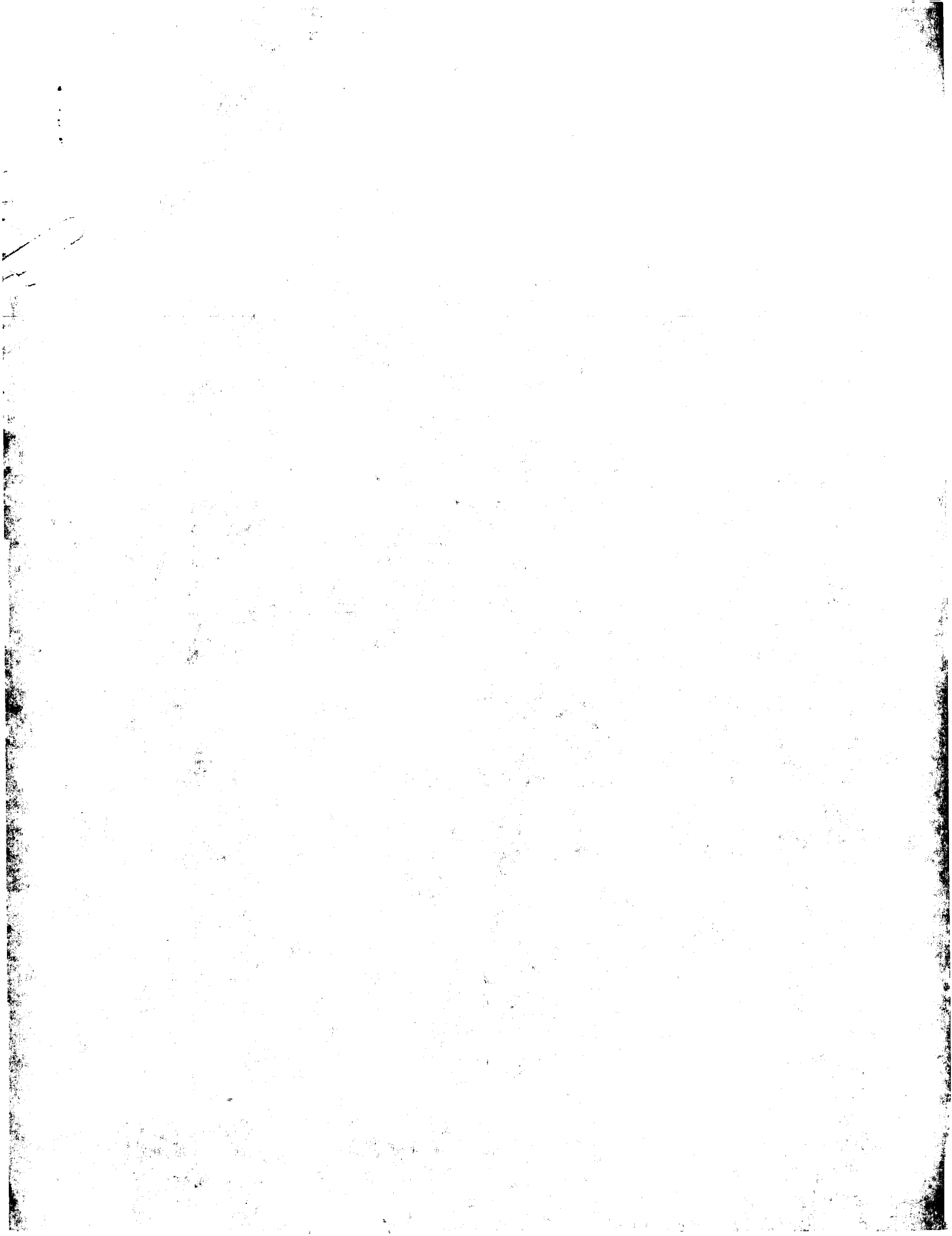
RESULT 15

Q9XVZ8 PRELIMINARY; PRT; 84 AA.
AC Q9XVZ8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Y54G11A.11 PROTEIN.
GN Y54G11A.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wallis J.M.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL034488; CAA22454.1; -
SO SEQUENCE 84 AA; 9667 MW; E996ECFC218D227C CRC64;

Query Match 100.0%; Score 20; DB 5; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKMD 4
||||
DB 37 VKMD 40

Search completed: October 30, 2002, 12:30:15
Job time : 3.24816 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:20:50 ; Search time 5.60688 Seconds
(without alignments)
277.344 Million cell updates/sec

Title: US-09-724-571-72

Perfect score: 63
Sequence: 1 KTEEISEVNVAAEF 14

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	96.8	14	21	AA07888
2	60	95.2	19	22	AAE0612
3	54	85.7	18	22	AAE0069
4	54	85.7	19	22	AAE0069
5	54	85.7	20	21	AAE0611
6	54	85.7	20	21	AAE0611
7	54	85.7	20	18	AA08360
8	54	85.7	30	20	AAI33754
9	54	85.7	30	21	AA07895
10	54	85.7	30	22	AA07264
11	54	85.7	32	17	AA04402

12	54	85.7	32	17	AA04403	Mouse amyloid prec
13	54	85.7	32	17	AA04401	Mouse amyloid prec
14	54	85.7	33	18	AA08359	Beta-secretase sub
15	54	85.7	33	20	AAI33753	Synthetic oligopep
16	54	85.7	33	21	AA07892	Substrate for beta
17	54	85.7	39	21	AAE0611	Peptide 17-16'/SW,
18	54	85.7	42	18	AA08350	Beta-APP alpha-sec
19	54	85.7	42	20	AAI33752	Wild type APP beta
20	54	85.7	42	22	AA07262	Synthetic oligopep
21	54	85.7	58	20	AA08001	Peptide 26-16'/SW,
22	54	85.7	115	20	AA07997	Swedish-FAD APP714
23	54	85.7	115	20	AA08000	Swedish-FAD APP po
24	54	85.7	261	22	AAE0610	Chimeric cassette
25	54	85.7	506	19	AAE0610	Maltose binding pr
26	54	85.7	506	20	AAI33742	MBP-APP (SW192) fu
27	54	85.7	506	22	AA07258	MBP-APP C-125 (Swe
28	54	85.7	695	18	AA019504	APP695 mutant A-be
29	54	85.7	695	21	AA08435	APP695 mutant A-be
30	54	85.7	695	21	AA08435	Human APP695-sw va
31	54	85.7	695	22	AAE0633	Human amyloid prot
32	54	85.7	695	22	AAE06863	Human amyloid prec
33	54	85.7	695	22	AAU06607	Human amyloid prec
34	54	85.7	695	22	AAU07206	Human beta-amyloid
35	54	85.7	695	22	AAE02585	Human amyloid prec
36	54	85.7	695	22	AAE02585	Human APPSW-KK aml
37	54	85.7	697	21	AAE10636	Human amyloid prot
38	54	85.7	697	22	AAE06866	Human amyloid prec
39	54	85.7	697	22	AAU06610	Human amyloid prec
40	54	85.7	697	22	AAU07209	Human beta-amyloid
41	54	85.7	697	22	AAE02588	Human amyloid prec
42	54	85.7	751	18	AA019505	APP751 mutant A-be
43	54	85.7	751	18	AA019492	APP751 mutant A-be
44	54	85.7	770	18	AA019506	APP770 mutant A-be
45	54	85.7				

ALIGNMENTS

RESULT 1
ID AAB07888 standard; peptide; 14 AA.
XX AAB07888;
AC AAB07888;
XX AAB07888;
DT 14-NOV-2000 (first entry)
XX A peptide fragment derived from beta-amyloid precursor protein.
DE Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor.
XX Homo sapiens.
OS WO200047618-A2.
PN 17-AUG-2000.
XX 17-AUG-2000.
PD 10-FEB-2000; 2000WO-US03819.
XX 10-FEB-2000; 2000WO-US03819.
PR 10-FEB-1999; 99US-0119571.
XX 15-JUN-1999; 99US-0139172.
PA (ELAN-) ELAN PHARM INC.
XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX WPI; 2000-533011/48.
PT Purified beta-secretase protein used in assays to discover inhibitors
which can be used for the treatment of amyloidogenic diseases e.g.

```

PT Alzheimer's disease -
XX
PS Disclosure; Page 12; 121pp; English.
XX
CC The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or improve cognitive
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a peptide derived from beta-amyloid
CC precursor protein
XX
SQ Sequence 14 AA:
Query Match 96.8%; Score 61; DB 21; Length 14;
Best Local Similarity 92.9%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTEEISEVNVXVAEF 14
   |||||||
DB 1 KTEEISEVNVXVAEF 14

RESULT 2
AAE00612
ID AAE00612 standard; peptide; 19 AA.
XX
XX AAE00612;
XX
DT 02-JUL-2001 (first entry)
XX
DE PCBC1 APP beta-secretase cleavage site (D/I) mutant.
XX
KM Amyloid precursor protein; APP; caspase; beta-secretase; cleavage site;
KM cysteine protease; apoptosis; caspase expression cassette; metastasis;
KM tumour; cathepsin B; urokinase; proliferation; gene therapy;
KM Interdomain linker; Alzheimer's disease; mutant; muteln.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 12 /note= "Wild-type Asp substituted by Ile"
FT
PN WO200129232-A2.
XX
PD 26-APR-2001.
XX
PF 19-OCT-2000; 2000WO-US28941.
XX
PR 20-OCT-1999; 99US-0160559.
PR 14-AUG-2000; 2000US-0225564.
XX
PA (SCIO-) SCIOS INC.
XX
PI Cordell B, Li Y;
XX
PI Cordell B, Li Y;
XX
DR WPI; 2001-290920/30.
XX
PT Novel fusion polypeptide comprising first and second caspase subunit
PT separated by cleavage site not associated in nature with caspase
PT subunit, useful for cloning gene encoding enzymes involved in
PT proteolytic cleavage
XX
PS Disclosure; Fig 29A; 116pp; English.
XX
CC The present amino acid sequence is a PCBC1 plasmid amyloid precursor
CC protein (APP) beta-secretase cleavage site (AAE00611) (D/I) mutant. APP
CC sequence containing this mutation is less susceptible to beta-secretase

```

```

CC cleavage. APP beta-secretase cleavage site is used to construct an
CC artificially engineered chimeric cassette comprising human caspase-3 with
CC interdomain linker replaced by Swedish mutant beta-secretase cleavage
CC site. This modified caspase-3 plays a pivotal role in Alzheimer's
CC disease. Caspases are a family of cysteine proteases, that participate
CC in the initiation and execution of apoptosis.
CC The present invention relates to a method for functional cloning of genes
CC encoding proteins or enzymes involved in proteolytic cleavage. The
CC invention is based on the use of caspase expression cassettes comprising
CC the coding sequence of a proteolytic cleavage site flanked by sequences
CC encoding two caspase subunits. A fusion polypeptide comprising a first
CC and a second caspase subunit, separated by a cleavage site not associated
CC in nature, is useful for cloning gene encoding enzymes involved in
CC proteolytic cleavage. An expression cassette containing fusion
CC polypeptide is used to identify a mutant cell line deficient in an
CC enzyme of interest and is also useful for diagnosis and suppression of
CC proliferation or metastases of a tumour cell characterised by
CC overexpression of a polypeptide (e.g. Cathepsin B or urokinase,
CC selectively expressed in the tumour cells). DNA encoding fusion
CC polypeptide is used in gene therapy.
XX
SQ Sequence 19 AA:
Query Match 95.2%; Score 60; DB 22; Length 19;
Best Local Similarity 85.7%; Pred. No. 0.00015;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTEEISEVNVXVAEF 14
   |||||||
DB 2 KTEEISEVNVXVAEF 15

RESULT 3
AAE00609
ID AAE00609 standard; peptide; 18 AA.
XX
XX AAE00609;
XX
DT 02-JUL-2001 (first entry)
XX
DE Beta-amyloid precursor protein beta-secretase cleavage site mutant.
XX
KM Beta-amyloid precursor protein; beta-APP; caspase; beta-secretase;
KM cysteine protease; apoptosis; caspase expression cassette; metastasis;
KM tumour; cathepsin B; urokinase; proliferation; gene therapy;
KM Interdomain linker; cleavage site; Alzheimer's disease; mutant; muteln.
XX
OS Unidentified.
XX
PN WO200129232-A2.
XX
PD 26-APR-2001.
XX
PF 19-OCT-2000; 2000WO-US28941.
XX
PR 20-OCT-1999; 99US-0160559.
PR 14-AUG-2000; 2000US-0225564.
XX
PA (SCIO-) SCIOS INC.
XX
PI Cordell B, Li Y;
XX
PI Cordell B, Li Y;
XX
DR WPI; 2001-290920/30.
XX
PT Novel fusion polypeptide comprising first and second caspase subunit
PT separated by cleavage site not associated in nature with caspase
PT subunit, useful for cloning gene encoding enzymes involved in
PT proteolytic cleavage
XX
PS Example 2; Page 26; 116pp; English.
XX
CC The present sequence is a beta-secretase cleavage site of beta-amyloid
CC precursor protein (beta-APP) mutant found in certain families of

```

Query Match	85.7%;	Score 54;	DB 22;	Length 18;
Best Local Similarity	85.7%;	Pred. NO. 0.0018;		
Matches 12;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0

RESULT 4	
AAE00611	
ID	AAE00611 standard; peptide; 19 AA
XX	

any other precursor protein (APP) beta-secretase cleavage site

interdomain linker; Alzheimer's disease.

WO200129232-A2.

19-OCT-2000; 2000WO-US28941.

20-OCT-1999; 99US-0160559.
14-AUG-2000; 2000US-0225564.

WPI; 2001-290920/30.

separated by cleavage site not associated in nature with caspase subunit, useful for cloning gene encoding enzymes involved in proteolytic cleavage -

Disclosure; Fig 28A; 116pp; English.

The present amino acid sequence is a beta-secretase cleavage site of an amyloid precursor protein (APP). This sequence is used to construct an artificially engineered chimeric cassette comprising human caspase-3

CC with in-domain linker replaced by swedish mutant beta-secretase
CC cleavage site. This modified caspase-3 plays a pivotal role in
CC Alzheimer's disease. Caspases are a family of cysteine proteases, that
CC participate in the initiation and execution of apoptosis.
CC The present invention relates to a method for functional cloning of genes
CC encoding proteins or enzymes involved in proteolytic cleavage. The
CC invention is based on the use of caspase expression cassettes comprising
CC the coding sequence of a proteolytic cleavage site flanked by sequences
CC encoding two caspase subunits. A fusion polypeptide comprising a first
CC and a second caspase subunit, separated by a cleavage site not associated
CC in nature, is useful for cloning gene encoding enzymes involved in
CC proteolytic cleavage. An expression cassette containing fusion
CC polypeptide is used to identify a mutant cell line deficient in an
CC enzyme of interest and is also useful for diagnosis and suppression of
CC proliferation or metastases of a tumour cell characterised by
CC overexpression of a polypeptide (e.g. cathepsin B or urokinase,
CC selectively expressed in the tumour cells). DNA encoding fusion
CC polypeptide is used in gene therapy.

Query Match	85.7%;	Score 54;	DB 22;	Length 19;
Best Local Similarity	85.7%;	Pred. No. 0.0019;		
Matches 12; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

RESULT 5	
AAV69714	
ID	AAV69714 standard; peptide; 20 AA
XX	

beta-APP alpha-secretase substrate [NL]-APP(-10,+10).

neurotrophic; neuroprotective; beta-amyloid precursor protein; metabolism
cleavage site; beta-secretase; neurodegenerative disease;
Alzheimer's disease.

PN W09964587-A1

04-JUN-1999; 99WO-FR01326

05-JUN-1998; 98FR-0007068
31-MAR-1999; 99US-0122599

(RHON) RHONE-POULENC RORER SA.

WPI; 2000-097537/08.

Example 3; Page 24; 44pp; French.

7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530

novel polypeptide 1697) to represent synthetic peptide substrates for a specifically the natural beta-amyloid precursor protein (BAP). Norma cleavage of the protein occurs between amino acids Met396-Asp597 and Val636-Ile637 (positions 4-5 and 44-45 of A β 169701). The novel

CC polypeptide is used to identify agents that interact specifically with
 CC it. These agents regulate metabolism of APP, particularly they slow down
 CC or reduce production of beta-amyloid, so can be used to treat
 CC neurodegenerative diseases, particularly Alzheimer's disease.

XX Sequence 20 AA:

Query Match 85.7%; Score 54; DB 21; Length 20;

Best Local Similarity 85.7%; Pred. No. 0.002; Mismatches 2; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTEEISEVNXVAEF 14
 |||||
 DB 1 KTEEISEVNDMAEF 14

RESULT 6

AAV69716
 ID AAV69716 standard; peptide: 20 AA.

XX AAV69716;

XX 11-APR-2000 (first entry)

XX Beta-APP alpha-secretase substrate [NM]-APP(-10,+10).

XX Nootropic; neuroprotective; beta-amyloid precursor protein; metabolism;

XX cleavage site; beta-secretase; neurodegenerative disease;

KW Alzheimer's disease.

XX Homo sapiens.

OS Synthetic.

XX MO9964587-A1.

XX 16-DEC-1999.

XX 04-JUN-1999; 99WO-FR01326.

XX 05-JUN-1998; 98FR-0007068.

XX 31-MAR-1999; 99US-0122599.

XX (RHON) RHONE-POULENC ROKER SA.

XX (UYPA-) UNIV CURIE PARIS VI P & M.

XX Rholam M, Munoz-Gimenez N, Moutouakil M, Cohen P, Bertrand P;

XX WPI; 2000-097537/08.

XX Polypeptide with beta-secretase activity, specific for wild-type

XX amyloid precursor protein, useful in treating Alzheimer's disease

XX Example 3; Page 24; 44pp; French.

XX Peptides AAV69702-669718 represent synthetic peptide substrates for a

XX novel polypeptide with beta-secretase activity that can cleave

XX specifically the natural beta-amyloid precursor protein (bAPP). Normal

XX cleavage of the protein occurs between amino acids Met956-Asp97 and

XX Val636-Ile637 (positions 4-5 and 44-45 of AAV69701). The novel

XX polypeptide is used to identify agents that interact specifically with

XX it. These agents regulate metabolism of APP, particularly they slow down

XX or reduce production of beta-amyloid, so can be used to treat

XX neurodegenerative diseases, particularly Alzheimer's disease.

RESULT 7
 ID AAM08360 standard; peptide: 30 AA.

XX AAM08360;

XX 05-SEP-1997 (first entry)

XX Beta-secretase substrate #2.

XX Beta-cleavage site; beta-amyloid precursor protein; APP; beta-secretase;

XX alpha-secretase; proteolytic cleavage; inhibitor; Alzheimer's disease.

XX Synthetic.

XX MO9640885-A2.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US09985.

XX 07-JUN-1995; 95US-0485152.

XX 07-JUN-1995; 95US-0480498.

XX (ATHE-) ATHENA NEUROSCIENCES INC.

XX Anderson JP, Chrysler SMS, Jacobson-croak KL, Keim PS;

XX Mcconlogue LC, Sima S, Tan H;

XX WPI; 1997-052304/05.

XX Beta-secretase which specifically cleaves beta-amyloid precursor

XX protein useful to screen for inhibitors useful in treatment of

XX Alzheimer's disease

XX Disclosure; Page 45; 92pp; English.

XX AAM08359-W08362 represent substrates for the enzyme of the invention.

XX The enzyme of the invention is beta-secretase, and specifically cleaves

XX beta-amyloid precursor protein (beta-APP). Normal processing of beta-APP

XX is thought to occur via cleavage between residues 16 and 17 of the

XX beta-amyloid peptide region by an alpha-secretase. Pathogenic processing

XX is thought to occur by beta-secretase cleavage of beta-APP. A method of

XX beta-secretase activity can be detected and measured using a method of

XX the invention, which detects at least one of the beta-secretase cleavage

XX products formed on cleavage. The method can be used to determine whether

XX a test substance inhibits proteolytic cleavage, by beta-secretase, of

XX beta-APP. Compounds effective to at least partially inhibit

XX beta-secretase activity can be used to inhibit cleavage of beta-APP in

XX cells or mammalian hosts. Isolation and purification of beta-secretase

XX will permit chemical modelling of a critical event in the pathology of

XX Alzheimer's disease.

XX Sequence 30 AA:

Query Match 85.7%; Score 54; DB 18; Length 30;

Best Local Similarity 85.7%; Pred. No. 0.0032; Mismatches 2; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTEEISEVNXVAEF 14
 |||||
 DB 17 KTEEISEVNDMAEF 30

RESULT 8

AAV33754
 ID AAV33754 standard; protein: 30 AA.

XX AAV33754;

XX 09-NOV-1999 (first entry)

```
XX Synthetic oligopeptide 26-4'SW.
DE
XX
XX Beta-secretase: beta-amyloid protein precursor; APP; Down's syndrome;
KW Alzheimer's disease; measure activity; cleavage site.
XX
XX Synthetic.
OS
XX
XX US5942400-A.
PN
XX
XX 24-AUG-1999.
PD
XX
XX 07-JUN-1996; 96US-0659984.
PE
XX
XX 07-JUN-1996; 96US-0659984.
PR
XX 07-JUN-1995; 95US-0480498.
PR
XX 07-JUN-1995; 95US-0485152.
XX
XX (ELAN-) ELAN PHARM INC.
PA
XX
XX
XX Anderson JP, Jacobson-Croak KL, Sinha S;
PI
XX WPI: 1999-517417/43.
DR
XX
XX A method for detecting human beta-secretase cleavage of polypeptides
PT useful for identifying beta-secretase inhibitors
XX
XX Examples; Column 30; 43pp; English.
PS
XX
XX Sequences AAY33752-Y33756 are synthetic oligopeptides used for measuring
CC the activity of beta-secretase (AAY33741). Beta-secretase is capable of
CC cleaving beta-amyloid protein precursor (APP) (AAY33742). These
CC synthetic peptides contain the cleavage site of APP. Beta-secretase and
CC APP are used in a method for detecting human beta-secretase cleavage of
CC polypeptides and for identifying beta-secretase inhibitors. Inhibition
CC of beta-secretase activity would be useful for chemical modeling of a
CC critical event in the pathology of Alzheimer's disease. Inhibitors of
CC beta-secretase would be useful for the prevention and treatment of
CC Alzheimer's disease and Down's Syndrome.
CC
XX
XX Sequence 30 AA;
SQ
XX
XX
XX Query Match 85.7%; Score 54; DB 20; Length 30;
XX Best Local Similarity 85.7%; Pred. No. 0.0032;
XX Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 KTEISEVNXVAEF 14
XX ||||||| |||
DB 17 KTEISEVNLDAEF 30
XX
XX
XX RESULT 9
XX AAB07895
XX ID AAB07895 standard; Peptide; 30 AA.
XX
XX AAB07895;
XX
XX 14-NOV-2000 (first entry)
DT
XX
XX Substrate for beta-secretase enzyme.
DE
XX
XX Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
XX FT Cleavage-site 26..27
XX
XX W0200047618-A2.
XX
XX 17-AUG-2000.
XX
XX
```

```
XX
XX 10-FEB-2000; 2000WO-US03819.
PF
XX
XX 10-FEB-1999; 99US-0119571.
XX
XX 15-JUN-1999; 99US-0139172.
PR
XX
XX (ELAN-) ELAN PHARM INC.
PA
XX
XX
XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
PI
XX WPI: 2000-533011/48.
DR
XX
XX Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
XX Alzheimer's disease -
XX
XX Example 4; Page 71; 121pp; English.
PS
XX
XX The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a peptide substrate used to test the
CC activity of beta-secretase enzyme.
CC
XX
XX Sequence 30 AA;
SQ
XX
XX
XX Query Match 85.7%; Score 54; DB 21; Length 30;
XX Best Local Similarity 85.7%; Pred. No. 0.0032;
XX Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 KTEISEVNXVAEF 14
XX ||||||| |||
DB 17 KTEISEVNLDAEF 30
XX
XX
XX RESULT 10
XX AAB47264
XX ID AAB47264 standard; Peptide; 30 AA.
XX
XX AAB47264;
XX
XX 18-JUL-2001 (first entry)
DT
XX
XX Peptide 26-4'SW, for used in beta-secretase assay.
DE
XX
XX Beta-secretase; isotype; beta-amyloid precursor protein; APP;
KW beta-amyloid peptide; beta-AP; Alzheimer's disease; Down's syndrome;
KW HCMWA-D; Swedish mutation; maltose binding protein; MBP.
XX
XX Synthetic.
OS
XX
XX US6221645-B1.
XX
XX 24-APR-2001.
PD
XX
XX 07-JUN-1996; 96US-0660531.
PE
XX
XX 07-JUN-1995; 95US-0480498.
PR
XX
XX (ELAN-) ELAN PHARM INC.
PA
XX
XX Chrysler SMS, Sinha S, Keim PS, Anderson JP, Tan H, McConlogue LC;
PI WPI: 2001-315578/33.
XX
XX Novel antibody that specifically binds native beta-secretase protein,
XX
```

PT	useful for raising anti-idiotypic antibodies and for detecting or
PT	diagnosing pathological conditions related to presence of respective
PT	antigens
XX	
PS	Example; Column 29: 42pp; English.
XX	
CC	The sequences given in ABA47262-67 represent synthetic peptides
CC	containing the cleavage sites derived from wild-type beta-amyloid
CC	precursor protein (APP). These peptides were used in assays utilising
CC	partially purified beta-secretase to identify beta-secretase inhibitors.
CC	Beta-secretase is thought to be responsible for the pathogenic
CC	processing of APP to form beta amyloid peptide (beta-AP) in beta-AP
CC	related conditions, e.g. Alzheimer's disease, Downs syndrome, HCHWA-D
CC	etc. Beta-secretase has a molecular weight of 260-300 kD and will bind
CC	to wheat germ agglutinin but not to concanavalin A. Beta-secretase
CC	will cleave both the wild type and the Swedish mutation of APP.
XX	
SQ	Sequence 30 AA;
QY	1 KTEETSEVNVXAEF 14
Db	17 KTEETSEVNVDAEF 30
RESULT 11	
AAW04402	Best Local Similarity 85.7%; Score 54; DB 22; Length 30;
ID	AAW04402 standard; Protein; 32 AA.
XX	
AC	AAW04402;
XX	
DT	21-JUL-1997 (first entry)
XX	
XX	Mouse amyloid precursor protein exon 16 mutant ST59.
DE	
XX	
KW	Exon 16; murine; mouse; amyloid; precursor; protein; APP;
KW	humanisation; homozygous; heterozygous; human; Abeta; Swedish;
KW	familial; Alzheimer's; disease; FAD; mutation; tool; model;
KW	elucidation; pathology; symptomatology; screen; inhibition;
KW	transgenic; mutant; ST59.
XX	
OS	Mus spp.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 13
FT	/note="corresponding codon GAA"
FT	Misc-difference 15
FT	/note="wild type Lys substituted with Asn"
FT	Misc-difference 16
FT	/note="wild type Met substituted with Leu"
FT	Misc-difference 21
FT	/note="wild type Gly substituted with Arg"
XX	
PN	W09634097-A1.
XX	
PD	31-OCT-1996.
XX	
PF	26-APR-1996; 96MO-US05824.
XX	
PR	23-APR-1996; 96US-0636876.
PR	26-APR-1995; 95US-0429207.
XX	
PA	(CEPH-) CEPHALON INC.
PI	Hoffman EK, Reaume AG, Scott RW, Siman R, Trusko SP,
XX	
DR	WPI, 1996-497629/49.
DR	N-PSDB; AAT38667.
XX	
PT	transgenic mice with humanised amyloid precursor protein gene -

PT	having at least 1 Swedish FAD mutation, useful as tools or models to
PT	elucidate role of human A-beta in Alzheimer's disease
XX	
PS	Example 2; Fig 12; 123pp; English.
XX	
CC	The present sequence is the protein encoded by a partial sequence
CC	from exon 16 of the mouse amyloid precursor protein (APP) gene,
CC	into which 4 humanising base pair changes have been introduced. The
CC	exon was then used in the preparation of mice homozygous or
CC	heterozygous for a targeted APP encoding gene, comprising a human
CC	Abeta peptide encoding sequence in place of the endogenous murine
CC	sequence, and at least 1 Swedish Familial Alzheimer's disease (FAD)
CC	mutation. The mice can be used as tools, or models to elucidate the
CC	role of human Abeta in AD pathology and symptomatology. They can
CC	also be used to screen chemical compounds for the ability to
CC	inhibit <i>in vivo</i> processing of APP, to yield the human Abeta peptide
CC	by administering the chemical compounds to a mouse and measuring
CC	the relative amounts of amyloidogenic and nonamyloidogenic
CC	processing of APP in a sample from the mouse at an appropriate
CC	interval after administration of the chemical compounds.
XX	
SO	Sequence 32 AA:
QY	Query Match 85.7%; Score 54; DB 17; Length 32;
	Best Local Similarity 85.7%; Pred. No. 0.0035;
	Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
DB	7 KTEEISEVNXVAEF 14
	7 KTEEISEVNLDAEF 20
RESULT 12	
AAM04403	
ID	AAM04403 standard; Protein; 32 AA.
XX	
AC	AAM04403;
XX	
DT	21-JUL-1997 (first entry)
XX	
DE	Mouse amyloid precursor protein exon 16 mutant ST61.
XX	
KW	Exon 16; murine; mouse; amyloid; precursor; protein; APP;
KW	humanisation; homozygous; heterozygous; human; Abeta; Swedish;
KW	familial; Alzheimer's; disease; FAD; mutation; tool; model;
KW	elucidation; pathology; symptomatology; screen; inhibition;
KW	transgenic; mutant; ST61.
XX	
OS	Mus spp.
XX	
XX	
FT	Key Location/Qualifiers
FT	Misc-difference 13
FT	/note= "corresponding codon GAA"
FT	Misc-difference 15
FT	/note= "wild type Lys substituted with Asn"
FT	Misc-difference 16
FT	/note= "wild type Met substituted with Leu"
FT	Misc-difference 26
FT	/note= "wild type Phe substituted with Tyr"
FT	Misc-difference 29
FT	/note= "wild type Arg substituted with His"
XX	
PN	W09634097-A1.
XX	
PD	31-OCT-1996.
XX	
PF	26-APR-1996; 96MO-US05824.
XX	
XX	23-APR-1996; 96GS-0636876.
PR	26-APR-1995; 95US-0429207.
XX	
PA	(CEPH-) CEPHALON INC.
XX	

PI Hoffman EK, Reaume AG, Scott RW, Siman R, Trusko SP;
XX
XX WPI: 1996-497629/49.
DR N-PSDB: AAT38668.
XX
PT Transgenic mice with humanised amyloid precursor protein gene -
PT having at least 1 Swedish FAD mutation, useful as tools or models to
PT elucidate role of human A-beta in Alzheimer's disease
XX
XX Example 2: Fig 12: 123pp: English.
XX
CC The present sequence is the protein encoded by a partial sequence
CC from exon 16 of the mouse amyloid precursor protein (APP) gene,
CC into which 6 humanising base pair changes have been introduced. The
CC exon was then used in the preparation of mice homozygous or
CC heterozygous for a targeted APP encoding gene, comprising a human
CC Abeta peptide encoding sequence in place of the endogenous murine
CC mutation. The mice can be used as tools, or models to elucidate the
CC role of human Abeta in AD pathology and symptomatology. They can
CC also be used to screen chemical compounds for the ability to
CC inhibit in vivo processing of APP, to yield the human Abeta peptide
CC by administering the chemical compounds to a mouse and measuring
CC the relative amounts of amyloidogenic and nonamyloidogenic
CC processing of APP in a sample from the mouse at an appropriate
CC interval after administration of the chemical compounds.
XX
SQ Sequence 32 AA:

Query Match 85.7%; Score 54; DB 17; Length 32;
Best Local Similarity 85.7%; Pred. No. 0.0035;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTEEISEVNXVAEF 14
Db 7 KTEEISEVNLDAEF 20
||||||| |||
KTEEISEVNLDAEF 20

RESULT 13
AAW04401
ID AAW04401 standard; Protein: 32 AA.
XX
AC AAW04401;
XX
DT 21-JUL-1997 (first entry)
XX
DE Mouse amyloid precursor protein exon 16 mutant ST58.
XX
KW Exon 16: murine; mouse; amyloid; precursor; protein; APP;
KW humanisation; homozygous; heterozygous; human; Abeta; Swedish;
KW familial; Alzheimer's; disease; FAD; mutation; tool; model;
KW elucidation; pathology; symptomatology; screen; inhibition;
KW transgenic; mutant; ST58.
XX
XX Mus spp.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 13 /note= "corresponding codon GAA"
XX FT Misc-difference 15 /note= "wild type Lys substituted with Asn"
XX FT Misc-difference 16 /note= "wild type Met substituted with Leu"
XX FT /note= "wild type Met substituted with Leu"
XX PN W09634097-A1.
XX
XX 31-OCT-1996.
XX
XX 26-APR-1996; 96WO-US05824.
XX
XX 23-APR-1996; 96US-0636876.
XX 26-APR-1995; 95US-0429207.
XX

PA (CEPH-) CEPHALON INC.
XX
XX Hoffman EK, Reaume AG, Scott RW, Siman R, Trusko SP;
PI
XX WPI: 1996-497629/49.
DR N-PSDB: AAT38666.
XX
PT Transgenic mice with humanised amyloid precursor protein gene -
PT having at least 1 Swedish FAD mutation, useful as tools or models to
PT elucidate role of human A-beta in Alzheimer's disease
XX
XX Example 2: Fig 12: 123pp: English.
XX
CC The present sequence is the protein encoded by a partial sequence
CC from exon 16 of the mouse amyloid precursor protein (APP) gene,
CC into which 3 humanising base pair changes have been introduced. The
CC exon was then used in the preparation of mice homozygous or
CC heterozygous for a targeted APP encoding gene, comprising a human
CC Abeta peptide encoding sequence in place of the endogenous murine
CC mutation. The mice can be used as tools, or models to elucidate the
CC role of human Abeta in AD pathology and symptomatology. They can
CC also be used to screen chemical compounds for the ability to
CC inhibit in vivo processing of APP, to yield the human Abeta peptide
CC by administering the chemical compounds to a mouse and measuring
CC the relative amounts of amyloidogenic and nonamyloidogenic
CC processing of APP in a sample from the mouse at an appropriate
CC interval after administration of the chemical compounds.
XX
SQ Sequence 32 AA:

Query Match 85.7%; Score 54; DB 17; Length 32;
Best Local Similarity 85.7%; Pred. No. 0.0035;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTEEISEVNXVAEF 14
Db 7 KTEEISEVNLDAEF 20
||||||| |||
KTEEISEVNLDAEF 20

RESULT 14
AAW08359
ID AAW08359 standard; peptide: 33 AA.
XX
AC AAW08359;
XX
DT 05-SEP-1997 (first entry)
XX
DE Beta-secretase substrate #1.
XX
XX Beta-cleavage site; beta amyloid precursor protein; APP; beta-secretase;
KW alpha-secretase; proteolytic cleavage; inhibitor; Alzheimer's disease.
XX
XX Synthetic.
XX
XX W09640885-A2.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US09985.
XX
XX 07-JUN-1995; 95US-0485152.
XX 07-JUN-1995; 95US-0480498.
XX
XX (ATHE-) ATHENA NEUROSCIENCES INC.
XX
XX Anderson JP, Chrysler SMS, Jacobson-croak KL, Keim PS;
PI Mcconlogue LC, Sinha S, Tan H;
PI WPI: 1997-052304/05.
XX
XX Beta-secretase which specifically cleaves beta-amyloid precursor
PT protein - useful to screen for inhibitors useful in treatment of

PT Alzheimer's disease
 XX
 PS Disclosure; Page 44; 92pp; English.
 XX
 CC AAM08359-W08362 represent substrates for the enzyme of the invention.
 CC beta-amyloid precursor protein (beta-APP). Normal processing of beta-APP
 CC is thought to occur via cleavage between residues 16 and 17 of the
 CC beta-amyloid peptide region by an alpha-secretase. Pathogenic processing
 CC is thought to occur by beta-secretase cleavage of beta-APP.
 CC Beta-secretase activity can be detected and measured using a method of
 CC the invention, which detects at least one of the beta-secretase cleavage
 CC products formed on cleavage. The method can be used to determine whether
 CC a test substance inhibits proteolytic cleavage, by beta-secretase, of
 CC beta-APP. Compounds effective to at least partially inhibit
 CC beta-secretase activity can be used to inhibit cleavage of beta-APP in
 CC cells or mammalian hosts. Isolation and purification of beta-secretase
 CC will permit chemical modelling of a critical event in the pathology of
 CC Alzheimer's disease.
 XX
 SQ Sequence 33 AA:
 Query Match 85.7%; Score 54; DB 18; Length 33;
 Best Local Similarity 85.7%; Pred. No. 0.0036;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 KTEIISEVNXVAEF 14
 |||||
 DB 8 KTEIISEVNLDAEF 21
 RESULT 15
 AAY33753
 ID AAY33753 standard; Protein: 33 AA.
 XX
 AC AAY33753;
 XX
 DT 09-NOV-1999 (first entry)
 XX
 DE Synthetic oligopeptide 17-16'SW.
 XX
 KM Beta-secretase; beta-amyloid protein precursor; APP; Down's syndrome;
 KM Alzheimer's disease; measure activity; cleavage site.
 XX
 OS Synthetic.
 XX
 PN US5942400-A.
 XX
 PD 24-AUG-1999.
 XX
 PF 07-JUN-1996; 96US-0659984.
 XX
 PR 07-JUN-1996; 96US-0659984.
 PR 07-JUN-1995; 95US-0480498.
 PR 07-JUN-1995; 95US-0485152.
 XX
 PA (ELAN-) ELAN PHARM INC.
 XX
 PI Anderson JP, Jacobson-Croak KL, Sinha S;
 DR WPI: 1999-517417/43.
 XX
 PT A method for detecting human beta-secretase cleavage of polypeptides
 PT useful for identifying beta-secretase inhibitors
 XX
 PS Examples; Column 30; 43pp; English.
 XX
 CC Sequences AAY33752-Y33756 are synthetic oligopeptides used for measuring
 CC the activity of beta-secretase (AAY33741). Beta-secretase is capable of
 CC cleaving beta-amyloid protein precursor (APP) (AAY33742). These
 CC synthetic peptides contain the cleavage site of APP. Beta-secretase and
 CC APP are used in a method for detecting human beta-secretase cleavage of
 CC polypeptides and for identifying beta-secretase inhibitors. Inhibition

CC of beta-secretase activity would be useful for chemical modelling of a
 CC critical event in the pathology of Alzheimer's disease. Inhibitors of
 CC beta-secretase would be useful for the prevention and treatment of
 CC Alzheimer's disease and Down's Syndrome.
 XX
 SQ Sequence 33 AA:
 Query Match 85.7%; Score 54; DB 20; Length 33;
 Best Local Similarity 85.7%; Pred. No. 0.0036;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 KTEIISEVNXVAEF 14
 |||||
 DB 8 KTEIISEVNLDAEF 21
 Search completed: October 30, 2002, 12:27:04
 Job time : 7.60688 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 30, 2002, 12:24:20 ; Search time 1.99509 Seconds
(without alignments)
171.400 Million cell updates/sec

Title: US-09-724-571-72

Perfect score: 63

Sequence: 1 KTEETSEVNXVAEF 14

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/PCRTUS_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	85.7	30	2	US-08-659-984A-17
2	54	85.7	30	4	US-08-660-531-17
3	54	85.7	33	2	US-08-659-984A-16
4	54	85.7	33	4	US-08-660-531-16
5	54	85.7	42	2	US-08-659-984A-15
6	54	85.7	42	4	US-08-660-531-15
7	54	85.7	506	2	US-08-659-984A-21
8	54	85.7	506	4	US-08-660-531-21
9	54	85.7	506	4	US-09-054-334-4
10	48	76.2	58	1	US-08-371-930-25
11	48	76.2	58	5	PCT-US94-01712-25
12	48	76.2	132	6	5187153-4
13	48	76.2	162	6	5220013-4
14	48	76.2	162	6	5223482-4
15	48	76.2	264	1	US-07-990-893-5
16	48	76.2	264	1	US-08-462-859A-9
17	48	76.2	487	1	US-08-122-659A-9
18	48	76.2	487	1	US-08-464-247A-9
19	48	76.2	487	1	US-08-464-248A-9
20	48	76.2	492	1	US-08-462-859A-7
21	48	76.2	492	1	US-08-123-659A-7
22	48	76.2	492	1	US-08-464-247A-7
23	48	76.2	492	1	US-08-464-248A-7
24	48	76.2	656	5	US-08-371-930-23
25	48	76.2	656	5	PCT-US94-01712-23
26	48	76.2	676	1	US-08-371-930-24
27	48	76.2	676	5	PCT-US94-01712-24

28	48	76.2	694	1	US-08-339-152A-18	Sequence 18, Appl
29	48	76.2	694	2	US-08-007-999B-5	Sequence 3, Appl
30	48	76.2	694	2	US-08-689-276A-5	Sequence 5, Appl
31	48	76.2	695	1	US-08-371-930-27	Sequence 2, Appl
32	48	76.2	695	1	US-08-123-702-2	Sequence 27, Appl
33	48	76.2	695	1	US-08-339-152A-30	Sequence 30, Appl
34	48	76.2	695	2	US-08-104-165-1	Sequence 1, Appl
35	48	76.2	695	3	US-08-464-250-1	Sequence 1, Appl
36	48	76.2	695	4	US-08-464-250-1	Sequence 1, Appl
37	48	76.2	695	4	US-09-458-481B-4	Sequence 4, Appl
38	48	76.2	695	4	US-09-458-481B-5	Sequence 5, Appl
39	48	76.2	695	4	US-09-458-481B-6	Sequence 6, Appl
40	48	76.2	695	4	US-09-458-481B-7	Sequence 7, Appl
41	48	76.2	695	4	US-09-458-481B-8	Sequence 8, Appl
42	48	76.2	695	5	PCT-US94-01712-27	Sequence 27, Appl
43	48	76.2	695	6	5218100-2	Patent No. 5218100
44	48	76.2	751	1	US-08-123-702-4	Sequence 4, Appl
45	48	76.2	751	2	US-08-104-165-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-659-984A-17
; Sequence 17, Application US/08659984A
; Patent No. 5942400
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Sinha, Sukanto
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
TITLE OF INVENTION: Inhibition
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,984A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-659-984A-17
Query Match 85.7%; Score 54; DB 2; Length 30;
Best Local Similarity 85.7%; Pred. No. 0.00038;
Matches 12; Conservative 0; Mismatches 2; Gaps 0;

OY 1 KTEISEVNXVAEF 14
| | | | | | | | | |
DB 17 KTEISEVNLDAEF 30

RESULT 2

US-08-660-531-17
; Sequence 17, Application US/08660531
; Patent No. 6221645
; GENERAL INFORMATION:
; APPLICANT: Chrysler, Susanna M.S.
; APPLICANT: Sinha, Sukanto
; APPLICANT: Keim, Pamela S.
; APPLICANT: Anderson, John P.
; TITLE OF INVENTION: Beta-Secretase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Ctr., 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,531
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,498
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-002210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-660-531-17

Query Match 85.7%; Score 54; DB 4; Length 30;
Best Local Similarity 85.7%; Pred. No. 0.00038; Indels 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 2;

OY 1 KTEISEVNXVAEF 14
| | | | | | | | | |
DB 17 KTEISEVNLDAEF 30

RESULT 3
US-08-659-984A-16
; Sequence 16, Application US/08659984A
; Patent No. 5942400

; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Sinha, Sukanto
; APPLICANT: Jacobson-Croak, Kirsten L.
; TITLE OF INVENTION: Assays for Detecting Beta-Secretase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Ctr., 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,984A
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,152
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-002810US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-659-984A-16

Query Match 85.7%; Score 54; DB 2; Length 33;
Best Local Similarity 85.7%; Pred. No. 0.00042;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTEISEVNXVAEF 14
| | | | | | | | | |
DB 8 KTEISEVNLDAEF 21

RESULT 4
US-08-660-531-16
; Sequence 16, Application US/08660531
; Patent No. 6221645

; GENERAL INFORMATION:
; APPLICANT: Chrysler, Susanna M.S.
; APPLICANT: Sinha, Sukanto
; APPLICANT: Keim, Pamela S.
; APPLICANT: Anderson, John P.
; TITLE OF INVENTION: Beta-Secretase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Ctr., 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,531
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,498
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:

NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-660-531-16

Query Match
Best Local Similarity 85.7%; Score 54; DB 4; Length 33;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
|||||
DB 8 KTEISEVNLDAEF 21

RESULT 5
US-08-659-984A-15
Sequence 15, Application US/08659984A
Patent No. 5942400
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Jacobson-Croak, Kirsten L.
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,984A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-659-984A-15

Query Match
Best Local Similarity 85.7%; Score 54; DB 2; Length 42;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
|||||
DB 17 KTEISEVNLDAEF 30

RESULT 6
US-08-660-531-15
Sequence 15, Application US/08660531
Patent No. 6221645
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Keim, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,531
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,498
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-660-531-15

Query Match
Best Local Similarity 85.7%; Score 54; DB 4; Length 42;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
|||||
DB 17 KTEISEVNLDAEF 30

RESULT 7
US-08-659-984A-21
Sequence 21, Application US/08659984A
Patent No. 5942400
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Jacobson-Croak, Kirsten L.
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,984A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2422
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-984A-21

Query Match 85.7%; Score 54; DB 2; Length 506;
Best Local Similarity 85.7%; Pred. No. 0.0096;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTEIISEVNXVAEF 14
DB 398 KTEIISEVNLDAEF 411

RESULT 8
US-08-660-531-21
Sequence 21, Application US/08660531
Patent No. 6221645
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Keim, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,531
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,498
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:

NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-660-531-21

Query Match 85.7%; Score 54; DB 4; Length 506;
Best Local Similarity 85.7%; Pred. No. 0.0096;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTEIISEVNXVAEF 14
DB 398 KTEIISEVNLDAEF 411

RESULT 9
US-09-054-334-4
Sequence 4, Application US/09054334
Patent No. 6329163
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Jacobson-Croak, Kirsten L.
APPLICANT: Sinha, Sukanto
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
TITLE OF INVENTION: Inhibition
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,334
FILING DATE: 02-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 015270-002820US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-054-334-4

Query Match 85.7%; Score 54; DB 4; Length 506;
Best Local Similarity 85.7%; Pred. No. 0.0096;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTEIISEVNXVAEF 14

Db 398 KTEISEVNMDAEF 411

RESULT 10

US-08-371-930-25
; Sequence 25, Application US/08371930
; Patent No. 5578451

GENERAL INFORMATION:

APPLICANT: Nishimoto, Ikuro

TITLE OF INVENTION: ALZHEIMER'S DISEASE THERAPEUTICS

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX

OPERATING SYSTEM: MS-DOS (Version 5.0)

SOFTWARE: WordPerfect (Version 5.1)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/371,930

FILING DATE:

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/019,208

FILING DATE: February 18, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 00786/154001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 58

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-371-930-25

Query Match 76.2%; Score 48; DB 1; Length 58;
Best Local Similarity 78.6%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNMVAAEF 14
Db 37 KTEISEVNMVAAEF 50

RESULT 11

PCT-US94-01712-25
; Sequence 25, Application PC/US9401712
; GENERAL INFORMATION:

APPLICANT: Nishimoto, Ikuro

TITLE OF INVENTION: ALZHEIMER'S DISEASE THERAPEUTICS

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX

OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01712
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/019,208

FILING DATE: February 18, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 00786/154001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 58

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

PCT-US94-01712-25

Query Match 76.2%; Score 48; DB 5; Length 58;
Best Local Similarity 78.6%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNMVAAEF 14
Db 37 KTEISEVNMVAAEF 50

RESULT 12

5187153-4

Patent No. 5187153

APPLICANT: CORDELL, BARBARA; SCHILLING, JAMES W.; KATUNUMA, NOBUHIKO

TITLE OF INVENTION: METHODS OF TREATMENT USING ALZHEIMER'S

AMYLOID POLYPEPTIDE DERIVATIVES

NUMBER OF SEQUENCES: 33

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/502,273

FILING DATE: 29-MAR-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 361,912

FILING DATE: 06-JUN-1989

APPLICATION NUMBER: 359,911

FILING DATE: 12-MAY-1989

APPLICATION NUMBER: 87,002

FILING DATE: 18-AUG-1987

APPLICATION NUMBER: 8,810

FILING DATE: 30-JAN-1987

APPLICATION NUMBER: 948,376

FILING DATE: 31-DEC-1986

APPLICATION NUMBER: 932,193

FILING DATE: 17-NOV-1986

SEQ ID NO: 4

LENGTH: 152

5187153-4

Query Match 76.2%; Score 48; DB 6; Length 152;
Best Local Similarity 78.6%; Pred. No. 0.035;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNMVAAEF 14
Db 64 KTEISEVNMVAAEF 77

RESULT 13

5220013-4

Patent No. 5220013

APPLICANT: PONTE, PHYLLIS A.;CORDELL, BARBARA
TITLE OF INVENTION: DNA SEQUENCE USEFUL FOR THE DETECTION
OF ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/444,118
FILING DATE: 30-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 87,002
FILING DATE: 18-AUG-1987
APPLICATION NUMBER: 8,810
FILING DATE: 30-JAN-1987
APPLICATION NUMBER: 948,376
FILING DATE: 31-DEC-1986
APPLICATION NUMBER: 932,193
FILING DATE: 17-NOV-1986
SEQ ID NO:4:
LENGTH: 162
5220013-4

Query Match 76.2%; Score 48; DB 6; Length 162;
Best Local Similarity 78.6%; Pred. No. 0.038;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
||||||| |||
DB 71 KTEISEVKMDAEF 84

RESULT 14
5223482-4
Patent No. 5223482
APPLICANT: SCHILLING, JAMES W.;PONTE, PHYLLIS A.;CORDELL,
BARBARA
TITLE OF INVENTION: RECOMBINANT ALZHEIMER'S PROTEASE
INHIBITORY AMYLOID PROTEIN AND METHOD OF USE
NUMBER OF SEQUENCES: 34
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/361,912
FILING DATE: 06-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 359,911
FILING DATE: 12-MAY-1989
APPLICATION NUMBER: 87,002
FILING DATE: 18-AUG-1987
APPLICATION NUMBER: 8,810
FILING DATE: 30-JAN-1987
APPLICATION NUMBER: 948,376
FILING DATE: 31-DEC-1986
APPLICATION NUMBER: 932,193
FILING DATE: 17-NOV-1986
SEQ ID NO:4:
LENGTH: 162
5223482-4

Query Match 76.2%; Score 48; DB 6; Length 162;
Best Local Similarity 78.6%; Pred. No. 0.038;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
||||||| |||
DB 71 KTEISEVKMDAEF 84

RESULT 15
US-07-990-893-5
Sequence 5, Application US/07990893
Patent No. 5547841
GENERAL INFORMATION:
APPLICANT: Marotta, Charles A.
APPLICANT: Zain, Sayeda
TITLE OF INVENTION: Genetic Sequences Coding For Alzheimer
TITLE OF INVENTION: Amyloid From Brain

NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/990,893
FILING DATE: 19921215
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0932.0250003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-990-893-5

Query Match 76.2%; Score 48; DB 1; Length 264;
Best Local Similarity 78.6%; Pred. No. 0.066;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
||||||| |||
DB 156 KTEISEVKMDAEF 169

Search completed: October 30, 2002, 12:32:29
Job time : 2.95509 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:23:30 ; Search time 2.68305 Seconds
(without alignments)
501.389 Million cell updates/sec

Title: US-09-724-571-72

Perfect score: 63
Sequence: 1 KTEISEVNXVAEF 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	48	76.2	82	P00438	Alzheimer's disease
2	48	76.2	695	A49795	Alzheimer's disease
3	48	76.2	695	A27485	Alzheimer's disease
4	48	76.2	695	S00550	Alzheimer's disease
5	48	76.2	770	Q8H0U4	Alzheimer's disease
6	42	66.7	747	JH0773	Alzheimer's disease
7	39	61.9	455	S50725	hypothetical prote
8	38	60.3	316	T23930	hypothetical prote
9	38	60.3	335	A24785	hypothetical prote
10	38	60.3	1354	T13930	tripeptidyl-peptid
11	37	58.7	1185	A42404	collagen adhesin -
12	36.5	57.9	587	T48382	auxin-regulated pr
13	36	57.1	512	G96555	55 kDa B regulator
14	36	57.1	513	S55889	protein phosphatas
15	36	57.1	650	T22002	hypothetical prote
16	36	57.1	837	A56681	hypothetical prote
17	36	57.1	837	JM0282	penicillin amidase
18	36	57.1	867	T05389	antigen 332 - mala
19	35	55.6	184	D72745	hypothetical prote
20	35	55.6	351	JU0056	flagellin - Serrat
21	35	55.6	433	C84619	probable serine ca
22	35	55.6	456	S19212	adenylosuccinate 1
23	35	55.6	456	C90829	adenylosuccinate 1
24	35	55.6	456	A85687	adenylosuccinate 1
25	35	55.6	504	AD1615	glycine betaine/ca
26	35	55.6	505	E95246	4-alpha-glucanotra
27	35	55.6	505	B98111	4-alpha-glucanotra
28	35	55.6	572	S72249	titthorax prote
29	35	55.6	731	T19721	hypothetical prote

30	35	55.6	802	2	S49252	penicillin amidase
31	35	55.6	802	2	I39665	penicillin amidase
32	34	54.0	136	2	C81436	biopolymer transpo
33	34	54.0	207	2	A47615	colanic acid capsu
34	34	54.0	225	2	E97244	colanic acid capsu
35	34	54.0	274	2	B96991	colanic acid capsu
36	34	54.0	294	2	D97693	citrate lyase, bet
37	34	54.0	294	2	AC2919	citrate lyase, bet
38	34	54.0	307	2	S38152	hypothetical prote
39	34	54.0	319	2	C97720	hypothetical prote
40	34	54.0	341	2	F71042	mip protein (impor
41	34	54.0	342	2	C75172	hypothetical prote
42	34	54.0	357	2	E71123	enolase related (e
43	34	54.0	366	2	B87446	probable sugar tra
44	34	54.0	412	2	C38351	hypothetical prote
45	34	54.0	415	2	T12974	phosphoprotein pho

ALIGNMENTS

RESULT 1

P00438 Alzheimer's disease amyloid A4 protein precursor - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995

C:Accession: P00438; C60045
R:Davidson, J.S.; West, R.L.; Kotikalapudi, P.; Maroun, L.E.

Biochem. Biophys. Res. Commun. 188, 905-911, 1992

A>Title: Sequence and methylation in the beta/A4 region of the rabbit amyloid precurs

A:Reference number: P00438; MUID:93075180

A:Accession: P00438
A:Molecule type: DNA

A:Residues: 1-82 <DAV>

A:Cross-references: GB:M83558; GB:M83657
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A>Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d

A:Reference number: A60045; MUID:92017079

A:Accession: C60045
A:Molecule type: mRNA

A:Residues: 12-68 <JOH>

A:Cross-references: EMBL:X56129
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinas

C:Keywords: alternative splicing; Alzheimer's disease; amyloid; Down's syndrome

Query Match 76.2% Score 48; DB 2; Length 82;

Best Local Similarity 76.6% Pred. No. 0.05; Mismatches 3; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KTEISEVNXVAEF 14

Db 7 KTEISEVXKMAEF 20

RESULT 2

A49795 Alzheimer's disease amyloid beta protein precursor - crab-eating macaque

C:Species: Macaca fascicularis (crab-eating macaque)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A49795

R:Podlinsky, M.B.; Tolan, D.R.; Selkoe, D.J.
Am. J. Pathol. 138, 1423-1435, 1991

A>Title: Homology of the amyloid beta protein precursor in monkey and human supports

A:Reference number: A49795; MUID:91273117

A:Accession: A49795

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-695 <POD>
A:Cross-references: GB:M58727; NID:g342062; PIDN:AAA36829.1; PID:g342063

C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinas

C:Keywords: alternative splicing

Query Match 76.2%; Score 48; DB 1; Length 695;
 Best Local Similarity 78.6%; Pred. No. 0.55;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
 |||||
 DB 587 KTEISEVKMDAEF 600

RESULT 3
 A27485
 Alzheimer's disease amyloid beta/A4 protein homolog precursor - mouse
 N:Alternate names: proteinase nexin II
 C:Species: Mus musculus (house mouse)
 C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 13-Aug-1999
 C:Accession: A27485; S19727; I49485
 R:Yamada, T.; Sasaki, H.; Furuya, H.; Miyata, T.; Goto, I.; Sakaki, Y.
 Biochem. Biophys. Res. Commun. 149, 665-671, 1987
 A:Title: Complementary DNA for the mouse homolog of the human amyloid beta protein precursor
 A:Reference number: A27485; MUID:88106489
 A:Accession: A27485
 A:Molecule type: mRNA
 A:Residues: 1-695 <YAM>
 A:Cross-references: GB:M18373; NID:g191568; PIDN:AAA37139.1; PID:g309085
 A:Experimental source: brain
 R:de Strooper, B.; van Leuven, F.; van den Berghe, H.
 Biochim. Biophys. Acta 1129, 141-143, 1991
 A:Title: The amyloid beta protein precursor or proteinase nexin II from mouse is closer
 A:Reference number: S19727; MUID:92096458
 A:Accession: S19727
 A:Molecule type: mRNA
 A:Residues: 1-210,'G','212-220','S','222-396','A','398-402','T','404-448','A','450-695 <STR>
 A:Cross-references: EMBL:X53979
 R:Iizumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, H.; Hattori, M.; Sakaki, Y.
 Gene 112, 189-195, 1992
 A:Title: Positive and negative regulatory elements for the expression of the Alzheimer's
 A:Reference number: I49485; MUID:92209998
 A:Accession: I49485
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-19 <RES>
 A:Cross-references: GB:D10603; NID:g220328; PIDN:BA01456.1; PID:g220329
 C:Genetics:
 A:Map position: 16C3
 C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
 C:Keywords: alternative splicing; amyloid; transmembrane protein

Query Match 76.2%; Score 48; DB 2; Length 695;
 Best Local Similarity 78.6%; Pred. No. 0.55;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
 |||||
 DB 587 KTEISEVKMDAEF 600

RESULT 4
 S00550
 Alzheimer's disease amyloid beta protein precursor - rat
 N:Alternate names: beta-A4 amyloid protein
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
 C:Accession: S00550; A41245; A39820; S46251
 R:Shivers, B.D.; Hilbich, C.; Multhaup, G.; Salbaum, M.; Beyreuther, K.; Seeburg, P.H.
 EMBO J. 7, 1365-1370, 1988
 A:Title: Alzheimer's disease amyloidogenic glycoprotein: expression pattern in rat brain
 A:Reference number: S00550; MUID:88312583
 A:Accession: S00550
 A:Molecule type: mRNA
 A:Residues: 1-695 <SHI>
 A:Cross-references: EMBL:X07648; NID:g55616; PIDN:CAA30488.1; PID:g55617
 R:Schubert, D.; Schroeder, R.; Lacomblere, M.; Salton, T.; Cole, G.
 Science 241, 223-226, 1988

A:Title: Amyloid beta protein precursor is possibly a heparan sulfate proteoglycan co
 A:Reference number: A41245; MUID:88264430
 A:Accession: A41245
 A:Molecule type: protein
 A:Residues: 18-37,'X','39-40','X','42-44 <SCH>
 A:Note: evidence for heparan sulfate attachment
 R:Hesse, L.; Beher, D.; Masters, C.L.; Multhaup, G.
 FEBS Lett. 349, 109-116, 1994
 A:Title: The beta-A4 amyloid precursor protein binding to copper.
 A:Reference number: S46251; MUID:94320627
 A:Contents: annotation; copper binding sites
 A:Note: rat peptides were isolated but not sequenced
 R:Potempska, A.; Styles, J.; Mehta, P.; Kim, K.S.; Miller, D.L.
 J. Biol. Chem. 266, 8464-8469, 1991
 A:Title: Purification and tissue level of the beta-amyloid peptide precursor of rat b
 A:Reference number: A39820; MUID:9121087
 A:Accession: A39820
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 18-32 <POT>
 A:Experimental source: brain
 C:Comment: Deposition of amyloid protein as neurofibrillary tangles and/or plaques is
 C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinas
 C:Keywords: alternative splicing; amyloid; glycoprotein; transmembrane protein
 F:625-648/Domain: transmembrane #status predicted <YAM>

Query Match 76.2%; Score 48; DB 2; Length 695;
 Best Local Similarity 78.6%; Pred. No. 0.55;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
 |||||
 DB 587 KTEISEVKMDAEF 600

RESULT 5
 OHRDUA
 Alzheimer's disease amyloid beta protein precursor [validated] - human
 N:Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor Xla inh
 N:Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascu
 protein precursor splice form APP(770)
 C:Species: Homo sapiens (man)
 C>Date: 30-Jun-1987 #sequence_revision 28-Jul-1995 #text_change 15-Sep-2000
 C:Accession: S02260; S05194; A32277; A33260; A33486; I39452; I39451; I39453; I59562;
 4688; A29583; A29302; A60805; J10038; S06121; A60355; A59011; A38384; S29076; S38252;
 R:Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayney, R.M.; Unterbeck, A.;
 Nucleic Acids Res. 17, 517-522, 1989
 A:Title: The Prec4(695) precursor protein of Alzheimer's disease A4 amyloid is encode
 A:Reference number: S02260; MUID:89128427
 A:Accession: S02260
 A:Molecule type: DNA
 A:Residues: 1-288,'V','365-770 <LEM1>
 A:Cross-references: EMBL:X13466
 A:Note: alternative splice form APP(695)
 R:Lemaire, H.G.
 submitted to the EMBL Data Library, November 1988
 A:Reference number: S05194
 A:Accession: S05194
 A:Molecule type: DNA
 A:Residues: 1-14,'VW','17-288','V','365-770 <LEM2>
 A:Cross-references: EMBL:X13466; NID:g35598; PIDN:CAA31830.1; PID:g871360
 A:Note: alternative splice form APP(695)
 R:La Faut, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.
 Biochem. Biophys. Res. Commun. 159, 297-304, 1989
 A:Title: Characterization of the 5'-end region and the first two exons of the beta-pr
 A:Reference number: A32277; MUID:89165870
 A:Accession: A32277
 A:Molecule type: DNA
 A:Residues: 1-75 <LAF>
 A:Cross-references: GB:M24546; GB:M24547; NID:g341202; PIDN:AC13654.1; PID:g516074
 R:Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little, S.P.
 Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989
 A:Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows simila

A:Reference number: A33260; MUID:89392030
A:Accession: A33260
A:Molecule type: DNA
A:Residues: 656-737 <JOH>
A:Cross-references: GB:M29270; NID:g178663; PIDN:AA51768.1; PID:g178665
R:Prelli, F.; Levy, E.; van Duijn, S.G.; Bots, G.T.A.M.; Luyendijk, W.; Frangione, B.
Biochem. Biophys. Res. Commun. 170, 301-307, 1990
A:Title: Expression of a normal and variant Alzheimer's beta-protein gene in amyloid of
A:Reference number: A35486; MUID:90321244
A:Accession: A35486
A:Molecule type: DNA
A:Residues: 672-710 <PRE1>
A:Note: 693-Gln was found in DNA isolated from HCMA-D patients
R:Yoshikuni, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sasaki, Y.
Gene 87, 257-263, 1990
A:Title: Genomic organization of the human amyloid beta-protein precursor gene.
A:Reference number: 139451; MUID:90236318
A:Accession: 139451
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EME
A:Molecule type: DNA
A:Residues: 1-770 <YOS1>
A:Cross-references: GB:M3112; NID:g178613; PIDN:AA59502.1; PID:g178616
A:Accession: 139451
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EME
A:Molecule type: DNA
A:Residues: 1-530, QWLMPTVPAFWKVGK <YOS2>
A:Cross-references: GB:M34875; NID:g178608; PIDN:AA59501.1; PID:g178615
R:Yoshikuni, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sasaki, Y.
Gene 102, 291-292, 1991
A:Reference number: A59020; MUID:91340168
A:Accession: A59020
A:Contents: annotation; erratum
A:Note: revised physical map for reference 139451
R:Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.; van Duine
Science 248, 1124-1126, 1990
A:Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrh
A:Reference number: 139453; MUID:90260663
A:Accession: 139453
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 656-737 <LEV>
A:Cross-references: GB:M37896; NID:g178618; PIDN:AA5127.1; PID:g178620
A:Note: a mutation with 693-Gln is presented
R:Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.
Science 254, 97-99, 1991
A:Title: A mutation in the amyloid precursor protein associated with hereditary Alzheimer
A:Reference number: 139562; MUID:92022553
A:Accession: 139562
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 689-716, 'F', 718-737 <MGR>
A:Cross-references: GB:S57665; NID:g236720; PIDN:AA51991.1; PID:g236721
R:Kamato, K.; Orr, H.T.; Payami, H.; Wajsbom, E.M.; Alonso, M.E.; Palst, S.M.; Anderson,
ataksis, S.E.; Korenberg, J.R.; Sharma, V.; Kukull, W.; Larson, E.; Heston, L.L.; Martin,
Am. J. Hum. Genet. 51, 998-1014, 1992
A:Title: Linkage and mutational analysis of familial Alzheimer disease kindreds for the
A:Reference number: AA4017; MUID:93035397
A:Accession: AA4017
A:Molecule type: DNA
A:Residues: 687-692, 'G', 694-718 <KAM1>
A:Cross-references: GB:S45136; NID:g257379; PIDN:AA52646.1; PID:g257380
A:Experimental source: familial Alzheimer disease family 1IT
A:Note: sequence extracted from NCBI backbone (NCBI:115376)
R:Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.; Grzeschik, K.H.;
Nature 325, 733-736, 1987
A:Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface
A:Reference number: A03134; MUID:87144572

A:Accession: A03134
A:Molecule type: mRNA
A:Residues: 1-268, 'V', 365-770 <KAN>
A:Cross-references: GB:Y00264; NID:g28525; PIDN:CAA68374.1; PID:g28526
A:Note: alternative splice form APP(695)
R:Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987
A:Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascula
A:Reference number: A29030; MUID:87231971
A:Accession: A29030
A:Molecule type: mRNA
A:Residues: 284-288, 'V', 365-646, 'E', 648-770 <ROB>
A:Cross-references: GB:M16765; NID:g178539; PIDN:AA51722.1; PID:g178540
R:Goldgaber, D.; Herman, M.I.; McBride, O.W.; Saffioti, U.; Gajdusek, D.C.
Science 235, 877-880, 1987
A:Title: Characterization and chromosomal localization of a cDNA encoding brain amylo
A:Reference number: AA7584; MUID:87120328
A:Accession: AA7584
A:Molecule type: mRNA
A:Residues: 674-756, 'S', 758-770 <GOL>
A:Cross-references: GB:M1533; NID:g178706; PIDN:AA35540.1; PID:g178707
A:Experimental source: brain
R:Tanzi, R.E.; Gusella, J.F.; Watkins, P.C.; Bruns, G.A.P.; St George-Hyslop, P.; Van
Science 235, 880-884, 1987
A:Title: Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near
A:Reference number: AA7585; MUID:87120329
A:Accession: AA7585
A:Molecule type: mRNA
A:Residues: 674-703 <TAN1>
A:Cross-references: GB:M15332; NID:g177957; PIDN:AA51564.1; PID:g177958
R:Dykes, T.; Weidemann, A.; Multhaup, G.; Salbaum, J.M.; Lemaire, H.G.; Kang, J.; Mue
EMBO J. 7, 949-957, 1988
A:Title: Identification, transmembrane orientation and biogenesis of the amyloid A4 p
A:Reference number: S02638; MUID:88296437
A:Accession: S02638
A:Molecule type: mRNA
A:Residues: 672-678 <DYR>
R:Tanzi, R.E.; McClatchey, A.I.; Lamertl, E.D.; Villa-Komaroff, L.; Gusella, J.F.; N
Nature 331, 528-530, 1988
A:Title: Protease inhibitor domain encoded by an amyloid protein precursor mRNA assoc
A:Reference number: S00707; MUID:88122640
A:Accession: S00707
A:Molecule type: mRNA
A:Residues: 286-344, 'I', 365-366 <TAN2>
A:Cross-references: EMBL:X06982; NID:g28817; PIDN:CAA30042.1; PID:g2929612
A:Experimental source: promyelocytic leukemia cell line HL60
R:Ponte, P.; Gonzalez-Demhilt, P.; Schilling, J.; Miller, J.; Hsu, D.; Greenberg, B.;
Nature 331, 525-527, 1988
A:Title: A new A4 amyloid mRNA contains a domain homologous to serine proteinase inh
A:Reference number: S00925; MUID:88122639
A:Accession: S00925
A:Molecule type: mRNA
A:Residues: 1-344, 'I', 365-770 <PO2>
A:Cross-references: GB:X06989; EMBL:Y00297; NID:g28720; PIDN:CAA30050.1; PID:g28721
A:Note: alternative splice form APP(751)
R:Kitaguchi, N.; Takahashi, Y.; Tokushima, Y.; Shiojiri, S.; Ito, H.
Nature 331, 530-532, 1988
A:Title: Novel precursor of Alzheimer's disease amyloid protein shows protease inhibi
A:Reference number: A38949; MUID:88122641
A:Accession: A38949
A:Molecule type: mRNA
A:Residues: 287-367 <KIT>
A:Cross-references: GB:X06981; NID:g28816; PIDN:CAA30041.1; PID:g2929611
A:Experimental source: glioblastoma cell line
A:Note: alternative splice form APP(770)
R:Vitek, M.P.; Rasool, C.G.; de Sauvage, F.; Vitek, S.M.; Bartus, R.T.; Beer, B.; Ash
Brain Res. Mol. Brain Res. 4, 121-131, 1988
A:Title: Absence of mutation in the beta-amyloid cDNAs cloned from the brains of thre
A:Reference number: A30320
A:Accession: A30320
A:Status: not compared with conceptual translation

A:Molecule type: mRNA
 A:Residues: 284-288, 'V', 365-770 <VIT1>
 A:Accession: B30320
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 122-288, 'V', 365-770 <VIT2>
 A:Accession: C30320
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 506-770 <VIT3>
 R:Zain, S.B., Salim, M., Chou, W.G., Sajdel-Sulowska, E.M., Majocha, R.E., Marotta, C.A.
 Proc. Natl. Acad. Sci. U.S.A. 85, 929-933, 1988
 A:Title: Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer disease br
 A:Reference number: A31087; MUID:88124954
 A:Accession: A31087
 A:Molecule type: mRNA
 A:Residues: 507-770 <ZAI>
 A:Cross-references: GB:M8734; NID:q178572; PIDN:AAA51726.1; PID:q178573
 A:Note: the authors translated the codon GAA for residue 599 as Gly, ACC for residue 603
 8 as Val, GTG for residue 609 as Asn, AAT for residue 610 as Gly, and GGT for residue 65
 A:Note: the cited Genbank accession number, J03594, is not in release 101.0
 R:Masters, C.L., Multhaup, G., Simms, G., Potgiesser, J., Martins, R.N., Beyreuther, K.

Query Match 76.2%; Score 48; DB 1; Length 770;
 Best Local Similarity 78.6%; Pred. No. 0.62;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNVXVAEF 14
 |||||
 DB 662 KTEISEVVKMDAEF 675

RESULT 6
 JH0773
 Alzheimer's disease amyloid beta protein precursor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 10-Jun-1993 #sequence_revision 10-Jun-1993 #text_change 13-Aug-1999
 C:Accession: JH0773
 R:Okado, H.; Okamoto, H.
 Biochem. Biophys. Res. Commun. 189, 1561-1568, 1992
 A:Title: A Xenopus homologue of the human beta-amyloid precursor protein: developmental
 A:Reference number: JH0773; MUID:93129227
 A:Accession: JH0773
 A:Molecule type: mRNA
 A:Residues: 1-747 <OKA>
 A:Cross-references: GB:S52417; NID:9263150; PIDN:AA24853.1; PID:9263151
 A:Experimental source: larva
 C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
 C:Keywords: alternative splicing; amyloid
 F:287-337/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

Query Match 66.7%; Score 42; DB 2; Length 747;
 Best Local Similarity 64.3%; Pred. No. 8.3;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNVXVAEF 14
 |||||
 DB 639 KTEISEVVKMDSEY 652

RESULT 7
 S50725
 hypothetical protein YNL206c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein N1346
 C:Species: Saccharomyces cerevisiae
 C>Date: 23-Aug-1995 #sequence_revision 19-Jan-1996 #text_change 21-Jul-2000
 C:Accession: S50725; S63163; S49863
 R:Joniaux, J.L.; Coster, F.; Purnelle, B.; Goffeau, A.
 Yeast 10, 1639-1645, 1994
 A:Title: A 21.7 kb DNA segment on the left arm of yeast chromosome XIV carries WH13, GCH
 A:Reference number: S50712; MUID:95242839
 A:Accession: S50725
 A:Status: nucleic acid sequence not shown

A:Molecule type: DNA
 A:Residues: 1455 <JON>
 A:Cross-references: EMBL:X78898; NID:g600045; PIDN:CAA55502.1; PID:g600059
 R:Coster, F.; Joniaux, J.L.; Goffeau, A.; Purnelle, B.; Van Dyck, L.
 submitted to the Protein Sequence Database, April 1996
 A:Reference number: S63151
 A:Accession: S63163
 A:Molecule type: DNA
 A:Residues: 1455 <COS>
 A:Cross-references: EMBL:Z71482; NID:g1302206; PID:e239612; PID:g1302207; MIPS:YNL206
 A:Experimental source: strain S288C
 C:Genetics:
 A:Map position: 14L

Query Match 61.9%; Score 39; DB 2; Length 455;
 Best Local Similarity 63.6%; Pred. No. 18;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTEISEVNVXV 11
 |||||
 DB 63 KTEISEINTI 73

RESULT 8
 T23930
 hypothetical protein R05D7.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T23930
 R:Dobson, R.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19819
 A:Accession: T23930
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-316 <WIL>
 A:Cross-references: EMBL:Z81105; PIDN:CA803217.1; GSPDB:GN00019; CESP:R05D7.3
 A:Experimental source: clone R05D7
 C:Genetics:
 A:Gene: CESP:R05D7.3
 A:Map position: 1
 A:Introns: 63/3; 97/3; 122/3; 195/2; 230/1

Query Match 60.3%; Score 38; DB 2; Length 316;
 Best Local Similarity 46.2%; Pred. No. 18;
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTEISEVNVXVAE 13
 :|||:|:|:|
 DB 283 ETEITEISEISE 295

RESULT 9
 A24785
 hypothetical protein 335 - slime mold (Dictyostelium discoideum) transposon DIRS-1
 C:Species: Dictyostelium discoideum
 C>Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 18-Jun-1993
 C:Accession: A24785
 R:Cappello, J.; Handelsman, K.; Lodish, H.F.
 Cell 43, 105-115, 1985
 A:Title: Sequence of Dictyostelium DIRS-1: an apparent retrotransposon with inverted
 A:Reference number: A94654; MUID:86079481
 A:Accession: A24785
 A:Molecule type: DNA
 A:Residues: 1-335 <CAP>

Query Match 60.3%; Score 38; DB 2; Length 335;
 Best Local Similarity 57.1%; Pred. No. 20;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KTEISEVNVXVAEF 14
 |:|:|:|:|
 DB 115 KKDISELNKVFNF 128

```

RESULT 10
T13930
tripeptidyl-peptidase II (EC 3.4.14.10) - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C:Accession: T13930
R:Renn, S.C.P.; Tomkinson, B.; Taghert, P.H.
J. Biol. Chem. 273, 19173-19182, 1998
A:Title: Characterization and cloning of tripeptidyl peptidase II from the fruit fly, D.
A:Reference number: 217814; MUID:98334655
A:Accession: T13930
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1354 <REN>
A:Cross-references: EMBL:AF035251; NID:g3387807; PID:g3387808; PIDN:AAC28563.1
C:Genetics:
A:Map position: 2
C:Keywords: dipeptidylpeptide hydrolase

Query Match
Best Local Similarity 60.3%; Score 38; DB 2; Length 1354;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 KEISEVNXVAEF 14
Db 929 EKISEANLVASF 940

RESULT 11
A42404
collagen adhesin - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 10-Nov-1995
C:Accession: A42404; S27655
R:Patt, J.M.; Jonsson, H.; Guss, B.; Switalski, L.M.; Wlberg, K.; Lindberg, M.; Hook, M.
J. Biol. Chem. 267, 4766-4772, 1992
A:Title: Molecular characterization and expression of a gene encoding a Staphylococcus a
A:Reference number: A42404; MUID:92165839
A:Contents: FDA 574
A:Accession: A42404
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1185 <PAT>
A:Cross-references: EMBL:M81736
A:Note: sequence extracted from NCBI backbone (NCBIP:83982)

Query Match
Best Local Similarity 58.7%; Score 37; DB 2; Length 1185;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNXVAEF 14
Db 118 KVEKLSVSGFAEF 131

RESULT 12
T48582
auxin-regulated protein GH3 homolog T31B5.170 - Arabidopsis thaliana
N:Alternate names: protein T31B5.170
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 18-Jul-2001
C:Accession: T48582
R:Bevan, M.; Hilbert, H.; Braum, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24490
A:Accession: T48582
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-587 <BEV>
A:Cross-references: EMBL:AL163491

```

```

A:Experimental source: cultivar Columbia; BAC clone T31B5
C:Genetics:
A:Map position: 5
A:Introns: 97/2; 131/2; 178/2; 400/1
A:Note: T31B5.170
C:Superfamily: soybean auxin-regulated protein GH3

Query Match
Best Local Similarity 57.9%; Score 36.5; DB 2; Length 587;
Matches 10; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

OY 1 KTEE---ISEVNXVAEF 14
Db 562 KTEEAVKILEANVASEF 578

RESULT 13
G96555
55 kDa B regulatory subunit of phosphatase 2A [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96555
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maltli, R.; Marzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: G96555
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <STO>
A:Cross-references: GB:AE05173; NID:g11094761; PIDN:AAG29694.1; GSPDB:GN00141
C:Genetics:
A:Gene: F19C24.10
A:Map position: 1

Query Match
Best Local Similarity 57.1%; Score 36; DB 2; Length 512;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KTEEISEVNXVAEF 14
Db 277 KMEDELSEVITSAEF 290

RESULT 14
S55889
protein phosphatase 2A B regulatory chain 55K - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S55889
R:Rundell, S.J.; Hartung, A.J.; Corum III, J.W.; O'Neill, M.
Plant Mol. Biol. 28, 257-266, 1995
A:Title: Characterization of a cDNA encoding the 55 kDa B regulatory subunit of Arabi
A:Reference number: S55889; MUID:95322588
A:Accession: S55889
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-513 <RUN>
A:Cross-references: GB:U18129; NID:g710329; PIDN:AAA86695.1; PID:g710330

Query Match
Best Local Similarity 57.1%; Score 36; DB 2; Length 513;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KTEEISEVNXVAEF 14

```

Db 278 KMEDISEVITSAEF 291

RESULT 15

T22002
hypothetical protein F39H11.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22002

R:White, S.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19500

A:Accession: T22002

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-650 <WILL>

A:Cross-references: EMBL:Z81079; PIDN:CAB03084.1; GSPDB:GN00019; CESP:F39H11.4

A:Experimental source: clone F39H11

C:Genetics:

A:Gene: CESP:F39H11.4

A:Map position: 1

A:introns: 109/3; 154/1; 357/1; 420/3; 464/3; 566/2; 594/1; 628/3

Query Match 57.1%; Score 36; DB 2; Length 650;

Best Local Similarity 42.9%; Pred. No. 99;

Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEIIEVNVXVAEF 14

||:| ||: ::|

Db 122 KQESSELQKISDF 135

Search completed: October 30, 2002, 12:31:27
Job time : 5.68305 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:10 ; Search time 1.20393 seconds

(without alignments)
450.253 Million cell updates/sec

Title: US-09-724-571-72

Perfect score: 63

Sequence: 1 KTEETSEVNVXAEF 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	76.2	751	1 A4_SAIISC	Q95241 salmirt sci
2	48	76.2	770	1 A4_HUMAN	P05067 homo sapien
3	48	76.2	770	1 A4_MOUSE	P12023 mus musculu
4	48	76.2	770	1 A4_RAT	P08592 rattus norv
5	39	61.9	455	1 YNU6_YEAST	P40161 saccharomyc
6	37	58.7	309	1 FSTL_FLABI	P52838 flaveria bi
7	37	58.7	351	1 CNA_STAU	O53654 staphylococ
8	35	55.6	456	1 FLIC_SERMA	P13713 scartitia ma
9	35	55.6	473	1 PUR8_ECOLI	P25739 escherichia
10	35	55.6	505	1 MDM2_XENLA	P38273 xenopus lae
11	35	55.6	505	1 MALO_STRPN	P29851 streptococc
12	35	55.6	802	1 PAC_ARTVI	P31956 arthrobacte
13	35	55.6	802	1 PAC_BACME	O60136 bacillus me
14	34	54.0	207	1 RCSA_KLEAE	P05338 klebsiella
15	34	54.0	307	1 YK55_YEAST	P36155 saccharomyc
16	34	54.0	413	1 ZABB_RABIT	Q00006 o serine/th
17	34	54.0	426	1 ZABA_PIG	Q29090 s serine/th
18	34	54.0	427	1 CISTY_ECOLI	P00891 escherichia
19	34	54.0	427	1 CISTY_SALTY	O68883 salmonella
20	34	54.0	443	1 ZABB_HUMAN	O00005 homo sapien
21	34	54.0	443	1 ZABB_PIG	P54614 sus scrofa
22	34	54.0	447	1 ZABB_RAT	P36877 r serine/th
23	34	54.0	447	1 ZABA_HUMAN	Q00007 h serine/th
24	34	54.0	447	1 ZABA_RAT	P36876 r serine/th
25	34	54.0	453	1 ZABD_RAT	P56932 r serine/th
26	34	54.0	499	1 ZABA_DROME	P36872 drosophila
27	34	54.0	620	1 Y870_METJA	O58280 methanococc
28	34	54.0	759	1 TEAL_YEAST	P47988 saccharomyc
29	34	54.0	967	1 CAP2_MAIZE	P51059 zea mays (m
30	34	54.0	1064	1 Y108_METJA	O60307 methanococc
31	34	54.0	1612	1 RRPO_ORSHVC	P86659 odontogloss
32	34	54.0	1613	1 VIT2_CAEEL	P05690 caenorhabdi
33	34	54.0	1616	1 VIT1_CAEEL	P55155 caenorhabdi

34	33	52.4	169	1	CHB1_AMPTPO	P05686 anthracia p
35	33	52.4	234	1	B10D_BACSH	P22818 bacillus sp
36	33	52.4	401	1	OD02_RICR	Q92094 rickettsia
37	33	52.4	423	1	CISTY_ACIAN	P20902 actinobact
38	33	52.4	452	1	YDDE_ECOLI	P38055 escherichia
39	33	52.4	901	1	PIP_LACLA	P49022 lactococcus
40	33	52.4	956	1	PODK_FLAAR	O42736 flaveria pr
41	33	52.4	1324	1	MSH6_ARATH	O04716 arabidopsis
42	32	50.8	94	1	STXE_BUTJU	P56637 buthocus ju
43	32	50.8	97	1	STXE_METUA	Q57802 methanococ
44	32	50.8	197	1	ENGB_ARCFU	O28943 archaeoglob
45	32	50.8	216	1	RCSB_ECOLI	P14374 escherichia

ALIGNMENTS

```

RESULT 1
A4_SAIISC STANDARD: PRT: 751 AA.
AC Q95241:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein precursor [Contains: Beta-
  amyloid protein (Beta-Ap) (A-beta)].
GN App.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N. A.
RC TISSUE=Liver and Kidney;
RX MEDLINE=96108492; PubMed=8532114;
RA Levy E., Amorim A., Frangione B., Walker L.C.;
RT "Beta-amyloid precursor protein gene in squirrel monkeys with
  cerebral amyloid angiopathy.";
RL Neurobiol. Aging 16:805-808(1995).
CC
CC -1- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
  INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
  G(O).
CC
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC
CC -1- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION
  WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC
  RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE
  NXYX MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF
  PHOSPHORYLATION (BY SIMILARITY).
CC
CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
CC
CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: S81024; AAD14347.1; -.
CC HSSP: P05067; IAP.
CC InterPro: IPR001868; A4_APP.
CC InterPro: IPR002223; Kunitz_BPTI.
CC Pfam: PF02177; A4_EXTRA; 1.
CC Pfam: PF00014; Kunitz_BPTI; 1.
CC PRINTS: PR00203; AMILOIDA4.
CC PRINTS: PR00759; BASICPTASE.
CC SMART: SM00131; KU; 1.
CC SMART: SM00319; A4_EXTRA; 1.
CC PROSITE: PS00319; A4_EXTRA; 1.
CC PROSITE: PS00320; A4_INTRA; 1.
CC PROSITE: PS00280; BPTI_KUNITZ_1; 1.
CC PROSITE: PS50279; BPTI_KUNITZ_2; 1.

```

KW Glycoprotein; Amyloid; Neurone; Transmembrane; Alternative splicing;
 KW Signal; Serine protease inhibitor.
 FT SIGNAL 1 17 BY SIMILARITY.
 FT CHAIN 18 751 A4 PROTEIN.
 FT CHAIN 653 695 BETA-AMYLOID PROTEIN (POTENTIAL).
 FT CHAIN 18 680 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 681 704 POTENTIAL.
 FT TRANSMEM 681 704 CYTOSOLASMIC (POTENTIAL).
 FT DOMAIN 705 751 BP1/KUNITZ INHIBITOR.
 FT SITE 287 345 CLATHRIN-BINDING (BY SIMILARITY).
 FT SITE 740 743 REACTIVE BINDING.
 FT ACT SITE 301 302 BY SIMILARITY.
 FT DISULFID 291 341 BY SIMILARITY.
 FT DISULFID 300 324 BY SIMILARITY.
 FT DISULFID 316 337 BY SIMILARITY.
 FT CARBOHYD 523 523 N-LINKED (GLCNAC. . .) (PROBABLE).
 FT CARBOHYD 552 552 N-LINKED (GLCNAC. . .) (PROBABLE).
 SQ SEQUENCE 751 AA; 84893 MW; 6C3E431089569049 CRC64;

Query Match 76.28; Score 48; DB 1; Length 751;
 Best Local Similarity 78.64; Pred. No. 0.35;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNXVAEF 14
 |||||
 DB 643 KTEISEVKMAEF 656

RESULT 2
 A4_HUMAN STANDARD; PRT; 770 AA.
 ID A4_HUMAN P05067; P09000; Q16011;
 AC 13-AUG-1987 (Rel. 05, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein precursor (Protease nexin-II)
 DE (PN-II) (APPI) [Contains: Beta-amyloid protein (beta-ApP) (A-beta)].
 GN APP OR A4 OR CVAP OR AD1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=87144572; PubMed=2881207;
 RA Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L.,
 RA Grzeschik K.-H., Multhaup G., Beyreuther K., Mueller-Hill B.;
 RT "The precursor of Alzheimer's disease amyloid A4 protein resembles a
 RT cell-surface receptor.";
 RL Nature 325:733-736(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88126359; PubMed=2893289;
 RA Ponte P., Gonzalez-Dewhilt P., Schilling J., Miller J., Hsu D.,
 RA Greenberg B., Davis K., Wallace W., Lieberburg I., Fuller F.,
 RA Cordell B.;
 RT "A new A4 amyloid mRNA contains a domain homologous to serine
 RT proteinase inhibitors.";
 RL Nature 331:525-527(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89128472; PubMed=2783775;
 RA Lemaire H.-G., Salbaum J.M., Multhaup G., Kang J., Bayney R.M.,
 RA Unterbeck A., Beyreuther K., Mueller-Hill B.;
 RT "The preA4(695) precursor protein of Alzheimer's disease A4 amyloid
 RT is encoded by 16 exons.";
 RL Nucleic Acids Res. 17:517-522(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97263807; PubMed=9108164;
 RA Hattori M., Tsukahara F., Furuhata Y., Tanahashi H., Hirose M.,
 RA -Saito M., Tsukuni S., Sakaki Y.;
 RT "A novel method for making nested deletions and its application for
 RT sequencing of a 300 kb region of human APP locus.";
 RL Nucleic Acids Res. 25:1802-1808(1997).
 RN [5]
 RP SEQUENCE OF 286-345 AND 365-366 FROM N.A.
 RX MEDLINE=88122640; PubMed=2893290;
 RA Tanzi R.E., McClatchey A.I., Lampetti E.D., Villa-Komaroff L.,
 RA Gusella J.F., Neve R.L.;
 RT "Protease inhibitor domain encoded by an amyloid protein precursor
 RT mRNA associated with Alzheimer's disease.";
 RL Nature 331:528-530(1988).
 RN [6]
 RP SEQUENCE OF 287-367 FROM N.A.
 RX MEDLINE=88122641; PubMed=2893291;
 RA Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.;
 RT "Novel precursor of Alzheimer's disease amyloid protein shows
 RT protease inhibitory activity.";
 RL Nature 331:530-532(1988).
 RN [7]
 RP SEQUENCE OF 284-289 AND 365-770 FROM N.A.
 RX MEDLINE=87231971; PubMed=3035574;
 RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;
 RT "Molecular cloning and characterization of a cDNA encoding the
 RT cerebrovascular and the neuritic plaque amyloid peptides.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).
 RN [8]
 RP SEQUENCE OF 507-770 FROM N.A.
 RX MEDLINE=88124954; PubMed=2893379;
 RA Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,
 RA Marotta C.A.;
 RT "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
 RT disease brain: coding and noncoding regions of the fetal precursor
 RT mRNA are expressed in the cortex.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).
 RN [9]
 RP SEQUENCE OF 672-681.
 RX MEDLINE=88035004; PubMed=3312495;
 RA Pardridge W.M., Vinters H.V., Yang J., Eisenberg J., Choi T.B.,
 RA Tourtelotte W.W., Huebner V., Shively J.E.;
 RT "Amyloid angiopathy of Alzheimer's disease: amino acid composition
 RT and partial sequence of a 4,200-dalton peptide isolated from cortical
 RT microvessels.";
 RL J. Neurochem. 49:1394-1401(1987).
 RN [10]
 RP SEQUENCE OF 739-770 FROM N.A.
 RX MEDLINE=90236318; PubMed=2110105;
 RA Yoshikawa S.-I., Sasaki H., Don-Ura K., Furuya H., Sakaki Y.;
 RT "Genomic organization of the human amyloid beta-protein precursor
 RT gene.";
 RL Gene 87:257-263(1990).
 RN [11]
 RP SEQUENCE OF 1-10 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89016647; PubMed=3140222;
 RA Schon E.A., Mita S., Sadlock J., Herbert J.;
 RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)
 RT encodes a 95-kDa polypeptide.";
 RL Nucleic Acids Res. 16:9351-9351(1988).
 RN [12]
 RP SEQUENCE OF 18-50.
 RX MEDLINE=87250462; PubMed=3597385;
 RA van Nostrand W.E., Cunningham D.D.;
 RT "Purification of protease nexin II from human fibroblasts.";
 RL J. Biol. Chem. 263:8508-8514(1987).
 RN [13]
 RP IDENTITY OF APP WITH NEXIN-II.
 RX MEDLINE=89384866; PubMed=2506449;
 RA Oltersdorf T., Fritz L.C., Schenk D.B., Lieberburg I., Dovey H.F.,
 RA Johnson-Wood K.L., Beattie E.C., Ward P.J., Blacher R.W.,
 RA Sinha S.;
 RT "The secreted form of the Alzheimer's amyloid precursor protein with
 RT the Kunitz domain is protease nexin-II.";
 RL Nature 341:144-147(1989).
 RN [14]

RP PROTEASE-SPECIFICITY OF INHIBITOR DOMAIN.
 RX MEDLINE-90211252; Pubmed-1969731;
 RA Kido H., Fukutomi A., Schilling J., Wang Y., Cordell B., Katunuma N.;
 RT "Protease-specificity of Kunitz inhibitor domain of Alzheimer's
 RT disease amyloid protein precursor.";
 RL Biochem. Biophys. Res. Commun. 167:716-721(1990).
 RN [15]
 RP COMPLEX WITH G(O).
 RX MEDLINE-9318965; Pubmed-8446172;
 RA Nishimoto I., Okamoto T., Matsura Y., Takahashi S., Okamoto T.,
 RA Miyama Y., Ogata E.;
 RT "Alzheimer amyloid protein precursor complexes with brain GTP-binding
 RT protein G(O).";
 RL Nature 362:75-79(1993).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 28-133.
 RX MEDLINE-99215582; Pubmed-10201399;
 RA Rosjohn J., Cappai R., Fell S.C., Henry A., McKinstry W.J.,
 RA Galatis D., Hesse L., Multhaup G., Beyreuther K., Masters C.L.,
 RA Parker M.W.;
 RT "Crystal structure of the N-terminal, growth factor-like domain of
 RT Alzheimer amyloid precursor protein.";
 RL Nat. Struct. Biol. 6:327-331(1999).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 287-344.
 RX MEDLINE-91104913; Pubmed-2125487;
 RA Hynes T.R., Randal M., Kennedy L.A., Eigenbrodt C., Kosiakof A.A.;
 RT "X-ray crystal structure of the protease inhibitor domain of
 RT Alzheimer's amyloid beta-protein precursor.";
 RL Biochemistry 29:10018-10022(1990).
 RN [18]
 RP STRUCTURE BY NMR OF 289-344.
 RX MEDLINE-92031488; Pubmed-1718421;
 RA Head S.L., Tilton R.F. Jr., Hammond L.S., Lee A., Bayney R.M.,
 RA Kamarek M.E., Ramabhadran T.V., Dreyer R.N., Davis G., Unterbeck A.,
 RA Tamburini P.P.;
 RT "Sequential NMR resonance assignment and structure determination of
 RT the Kunitz-type inhibitor domain of the Alzheimer's beta-amyloid
 RT precursor protein.";
 RL Biochemistry 30:10467-10478(1991).
 RN [19]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE-94281210; Pubmed-7516706;
 RA Talafous J., Marchowski K.J., Klopman G., Zagorski M.G.;
 RT "Solution structure of residues 1-28 of the amyloid beta-peptide.";
 RL Biochemistry 33:7788-7796(1994).
 RN [20]
 RP STRUCTURE BY NMR OF 696-706.
 RX MEDLINE-97128622; Pubmed-8973180;
 RA Kohno T., Kobayashi K., Maeda T., Sato K., Takashima A.;
 RT "Three-dimensional structures of the amyloid beta peptide (25-35) in
 RT membrane-mimicking environment.";
 RL Biochemistry 35:16094-16104(1996).
 RN [21]
 RP STRUCTURE BY NMR OF 672-711.
 RX MEDLINE-98359783; Pubmed-9693002;
 RA Coles M., Blacknell W., Watson A.A., Fairlie D.P., Craik D.J.;
 RT "Solution structure of amyloid beta-peptide(1-40) in a water-micelle
 RT environment. Is the membrane-spanning domain where we think it is?";
 RL Biochemistry 37:11064-11077(1998).
 RN [22]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE-20400066; Pubmed-10940222;
 RA Poulsen S.-A., Watson A.A., Craik D.J.;
 RT "Solution structures in aqueous SDS micelles of two amyloid beta
 RT peptides of Abeta(1-28) mutated at the alpha-secretase cleavage
 RT site.";
 RL J. Struct. Biol. 130:142-152(2000).
 RN [23]
 RP STRUCTURE BY NMR OF 681-706.
 RX MEDLINE-20400065; Pubmed-10940221;
 RA Zhang S., Iwata K., Lachemann M.J., Peng J.W., Li S., Stimson E.R.,
 RA Lu Y., Felix A.M., Maggio J.E., Lee J.P.;

RT "The Alzheimer's peptide a beta adopts a collapsed coil structure in
 RT water.";
 RL J. Struct. Biol. 130:130-141(2000).
 RN [24]
 RP SIGNAL SEQUENCE CLEAVAGE SITE, AND TOPOLOGY.
 RX MEDLINE-88296437; Pubmed-2900137;
 RA Dykes T., Weidemann A., Multhaup G., Salbaum J.M., Lemaire H.-G.,
 RA Kang J., Mueller-Hill B., Masters C.L., Beyreuther K.;
 RT "Identification, transmembrane orientation and biogenesis of the
 RT amyloid A4 precursor of Alzheimer's disease.";

Query Match 76.2%; Score 48; DB 1; Length 770;
 Best Local Similarity 78.6%; Pred. NO. 0.36;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEEISEVNXVAEF 14
 |||||
 DB 662 KTEEISEVXMDAEF 675

RESULT 3
 A4_MOUSE STANDARD; PRT; 770 AA.
 ID A4_MOUSE
 AC P12023;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein homolog precursor
 DE (Amyloidogenic glycoprotein) (AG).
 GN App.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId:10090;
 RN [1]
 RP SEQUENCE OF 1-289 AND 365-770 FROM N.A.
 RC STRAIN-BALB/C; TISSUE-Brain;
 RX MEDLINE-92096458; Pubmed-1756177;
 RA de Strooper B., van Leuven F., van den Bergh H.;
 RT "The amyloid beta protein precursor or proteinase nexin II from mouse
 RT is closer related to its human homolog than previously reported.";
 RL Biochim. Biophys. Acta 1129:141-143(1991).
 RN [2]
 RP SEQUENCE OF 1-289 AND 365-770 FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE-88106489; Pubmed-3322280;
 RA Yamada T., Sasaki H., Furuya H., Miyata T., Goto I., Sasaki Y.;
 RT "Complementary DNA for the mouse homolog of the human amyloid beta
 RT protein precursor.";
 RL Biochem. Biophys. Res. Commun. 149:665-671(1987).
 RN [3]
 RP REVISIONS.
 RA Yamada T.;
 RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 289-364 FROM N.A.
 RC STRAIN-CD-1; TISSUE-Placenta;
 RX MEDLINE-89345111; Pubmed-2569710;
 RA Fukuchi K., Martin G.M., Deeb S.S.;
 RT "Sequence of the protease inhibitor domain of the A4 amyloid protein
 RT precursor of Mus domesticus.";
 RL Nucleic Acids Res. 17:5396-5396(1989).
 RN [5]
 RP SEQUENCE OF 1-19 FROM N.A.
 RX MEDLINE-92209998; Pubmed-1555768;
 RA Izumi R., Yamada T., Yoshikai S.I., Sasaki H., Hattori M.,
 RA Sakai Y.;
 RT "Positive and negative regulatory elements for the expression of the
 RT Alzheimer's disease amyloid precursor-encoding gene in mouse.";
 RL Gene 112:189-195(1992).
 RN [6]
 RP SEQUENCE OF 281-380 FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE-Brain, and Kidney;

RX MEDLINE=89149813; PubMed=2493250.
 RA Yamada T., Sasaki H., Dohura K., Goto I., Sakaki Y.;
 RT "Structure and expression of the alternatively-spliced forms of mRNA
 RT for the mouse homolog of Alzheimer's disease amyloid beta protein
 RT precursor."
 RL Biochem. Biophys. Res. Commun. 158:906-912(1998).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
 CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: AAA(770) IS EXPRESSED IN KIDNEY, AAA(751) IS
 CC WIDELY EXPRESSED. AAA(695) IS EXPRESSED IN BRAIN, KIDNEY AND
 CC LIVER.
 CC -1- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION
 CC WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC
 CC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE
 CC NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF
 CC PHOSPHORYLATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X59379; -; NOT ANNOTATED_CDS.
 CC EMBL; M18373; AAA37139.1; -;
 CC EMBL; X15210; CAA33280.1; -;
 CC EMBL; D10603; BAA01456.1; -;
 CC EMBL; M24397; AAA39929.1; -;
 CC PIR; A27485; A27485.
 CC PIR; S04855; S04855.
 CC PIR; S19727; S19727.
 CC HSP; P05067; 10CM.
 CC MGD; MGI:88059; APP.
 CC InterPro: IPR001868; A4_APP.
 CC InterPro: IPR002223; Kunitz_BPTI.
 CC Pfam: PF02177; A4_EXTRA; 1.
 CC PRINTS; PR00203; AMYLOIDA4.
 CC PRINTS; PR00759; BASICPTASE.
 CC SMART; SM00006; A4_EXTRA; 1.
 CC SMART; SM00131; KU; 1.
 CC PROSITE; PS00319; A4_EXTRA; 1.
 CC PROSITE; PS00320; A4_INTRA; 1.
 CC PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 CC PROSITE; PS50279; BPTI_KUNITZ_2; 1.
 CC GlycoProtex; Amyloid; Neurone; Transmembrane; Signal;
 CC KW Alternative splicing; Serine protease inhibitor.
 CC FT SIGNAL 1 17
 CC CHAIN 18 770
 CC FT
 CC FT DOMAIN 18 699
 CC FT TRANSMEM 724 770
 CC FT DOMAIN 673 715
 CC FT DOMAIN 287 345
 CC FT SITE 759 762
 CC FT DISULFID 291 341
 CC FT DISULFID 300 324
 CC FT DISULFID 316 337
 CC FT CARBOHYD 542 542
 CC FT CARBOHYD 571 571
 CC FT VARSPLIC 289 289
 CC FT VARSPLIC 290 364
 CC FT VARSPLIC 346 380
 CC SEQUENCE 770 AA; 86752 MW; 26C5DDE890CAFA CRC64;
 Query Match 76.2%; Score 48; DB 1; Length 770;

Best Local Similarity 78.6%; Pred. No. 0.36;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 KTEISEVNVAAEF 14
 Db 662 KTEISEVNVKDAEF 675
 RESULT 4
 ID A4_RAT STANDARD; PRT; 770 AA.
 AC P08592;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein homolog precursor
 DE (Amyloidogenic glycoprotein) (Ab).
 GN APP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE OF 1-289 AND 365-770 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=88312583; PubMed=2900758;
 RA Shivers B.D., Hlilich C., Muthaup G., Salbaum J.M., Beyreuther K.,
 RA Seeburg P.H.;
 RT "Alzheimer's disease amyloidogenic glycoprotein: expression pattern
 RT in rat brain suggests a role in cell contact."
 RL EMBO J. 7:1365-1370(1988).
 RN [2]
 RP SEQUENCE OF 289-364 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89183625; PubMed=2648331;
 RA Kang J., Mueller-Hill B.;
 RT "The sequence of the two extra exons in rat preA4."
 RL Nucleic Acids Res. 17:2130-2130(1989).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
 CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION
 CC WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC
 CC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE
 CC NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF
 CC PHOSPHORYLATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X07648; CAA30488.1; -;
 CC EMBL; X14066; CAA32229.1; -;
 CC PIR; S00550; S00550.
 CC PIR; S03607; S03607.
 CC HSP; P05067; 1AAP.
 CC InterPro: IPR001868; A4_APP.
 CC InterPro: IPR002223; Kunitz_BPTI.
 CC Pfam: PF02177; A4_EXTRA; 1.
 CC Pfam: PF00014; Kunitz_BPTI; 1.
 CC PRINTS; PR00203; AMYLOIDA4.
 CC PRINTS; PR00759; BASICPTASE.
 CC SMART; SM00006; A4_EXTRA; 1.
 CC SMART; SM00131; KU; 1.
 CC PROSITE; PS00319; A4_EXTRA; 1.
 CC PROSITE; PS00320; A4_INTRA; 1.

Query Match Similarity	61.9%	Score	39	DB	1	Length	455
Best Local Similarity	63.6%	Pred. No.	9.2%				
Matches	7	Conservative	2	Mismatches	2	Indels	0
				Gaps			0

RA Patil J.M., Jonsson H., Guss B., Switalski L.M., Wiberg K.

```

RA Lindberg M., Hoeoek M.;
RT "Molecular characterization and expression of a gene encoding a
RT Staphylococcus aureus collagen adhesin."
RT J. Biol. Chem. 267:4766-4772(1992).
RN [2]
RA ERRATUM.
RA Patli J.M., Jonsson H., Guss B., Switalski L.M., Wiberg K.,
RA Lindberg M., Hoeoek M.;
RA J. Biol. Chem. 269:11672-11672(1994).
RN [3]
RP COLLAGEN-BINDING DOMAIN.
RC STRAIN-PDA 574;
RX MEDLINE=94032261; PubMed=8218209;
RA Patli J.M., Boles J.O., Hoeoek M.;
RT "Identification and biochemical characterization of the ligand
RT binding domain of the collagen adhesin from Staphylococcus aureus."
RL Biochemistry 32:11428-11435(1993).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 169-318.
RX MEDLINE=97475225; PubMed=9334749;
RA Symersky J., Patli J.M., Carson M., House-Pomoeo K., Teale M.,
RA Moore D., Jin L., Schneider A., Delucas L.J., Hoeoek M.,
RA Narayana S.V.L.;
RT "Structure of the collagen-binding domain from a Staphylococcus
RT aureus adhesin."
RL Nat. Struct. Biol. 4:833-838(1997).
CC -1- FUNCTION: MEDIATES ATTACHMENT OF STAPHYLOCOCCAL CELLS TO
CC COLLAGEN-CONTAINING SUBSTRATA.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Cell wall.
CC -1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC IN THE REGION OF THE MEMBRANE ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M81736; AAA20874.1; -.
DR PDB: 1AMX; 24-JUN-98.
DR InterPro: IPR001899; Gram_pos_anchor.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; FALSE_NEG.
KW Signal; Repeat; Transmembrane; Cell wall; 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 1183 POTENTIAL.
FT DOMAIN 30 1157 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1158 1177 MEMBRANE ANCHOR (POTENTIAL).
FT DOMAIN 1178 1183 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 151 318 COLLAGEN-BINDING.
FT DOMAIN 533 1093 3 X 187 AA APPROXIMATE TANDEM REPEATS.
FT DOMAIN 1093 1157 IYS/PRO-RICH (CELL WALL-SPANNING).
FT DOMAIN 1151 1156 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT PROTEINS.
FT REPEAT 533 719 B1.
FT REPEAT 720 906 B2.
FT REPEAT 907 1093 B3.
SQ SEQUENCE 1183 AA; 133066 MW; B6A1CC072E575D76 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 1183;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```

```

DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Flagellin.
GN FLIC OR FLAF OR HAG.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=274;
RX MEDLINE=89378766; PubMed=2673930;
RA Harshey R.M., Estepa G., Yanagi H.;
RT "Cloning and nucleotide sequence of a flagellin-coding gene (hag)
RT from Serratia marcescens 274."
RL Gene 79:1-8(1989)
CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC -1- FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M27219; AAA26556.1; -.
DR PIR: J00056; J00056.
DR InterPro: IPR001029; Flagellin_C.
DR InterPro: IPR001492; Flagellin_N.
DR Pfam: PF00700; Flagellin_C_1.
DR Pfam: PF00669; Flagellin_N_1.
DR PRINTS: PR00207; FLAGELLIN.
DR ProDom: PD000316; Flagellin_C_1.
KW Flagella.
SQ SEQUENCE 351 AA; 36865 MW; FDA4809AEB7D3CB8 CRC64;

Query Match 55.6%; Score 35; DB 1; Length 351;
Best Local Similarity 41.7%; Pred. No. 38;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

```

RESULT 8
ID *FLIC_SERMA STANDARD; PRT; 351 AA.
AC P13713;

```

```

RESULT 9
ID *PUB8_ECOLI STANDARD; PRT; 456 AA.
AC P25739;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylsuccinate lyase (EC 4.3.2.2) (Adenylsuccinase) (ASL).
GN PUB8 OR B1131.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-6.
RC STRAIN=K12;
RX MEDLINE=92104952; PubMed=1729205;
RA He B., Smith J.M., Zalkin H.;
RT "Escherichia coli purB gene: cloning, nucleotide sequence, and
RT regulation by purR."
RL J. Bacteriol. 174:130-136(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;

```

RA Green S.M., Drabble W.T.;
 RL Submitted (May-1991) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-97061202; PubMed-8905232;
 RA Oshima T., Alta H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiiuchi T.;
 RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
 corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 CC -1- CATALYTIC ACTIVITY: 1-(5'-PHOSPHORIBOSYL)-4-(N-SUCCINO-CARBOXYAMIDE)
 CC -5-AMINOIMIDAZOLE = FUMARATE + 5'-PHOSPHORIBOSYL-5-AMINO-4-
 CC IMIDAZOLECARBOXYAMIDE (ALSO CATALYZES: N6-(1,2-DICARBOXYETHYL)AMP =
 CC FUMARATE + AMP).
 CC -1- PATHWAY: EIGHT STEP IN DE NOVO PURINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE LYASE I FAMILY. ADENYLOSUCINATE LYASE
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M74924; AAA92731.1; -;
 DR EMBL: X59307; CAA41996.1; -;
 DR EMBL: AE000213; AAC74215.1; -;
 DR EMBL: D90748; BAA35953.1; -;
 DR EMBL: D90749; BAA35962.1; -;
 DR PIR: A43307; A43307.
 DR PIR: S19212; S19212.
 DR HSSP: P02633; 1BOD.
 DR Ecocore: EG1314; 1BOD.
 DR InterPro: IPR000362; Fumarate-lyase.
 DR Pfam: PF00206; lyase_1; 1.
 DR PROSITE: PS00163; FUMARATE_LYASES; 1.
 KW Purine biosynthesis; Lyase; Complete proteome.
 FT ACT_SITE 91 91 ACID (BY SIMILARITY).
 FT ACT_SITE 171 171 BASE (BY SIMILARITY).
 FT CONFLICT 145 145 P -> A (IN REF. 1).
 FT CONFLICT 154 154 I -> L (IN REF. 1).
 FT SEQUENCE 456 AA; 51542 MW; 8D1FA546B6795BC CRC64;
 Query Match 55.6%; Score 35; DB 1; Length 456;
 Best Local Similarity 50.0%; Pred. No. 50;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 KTEIISVYNXVAEF 14
 Db 103 KVAEIPELHAVSEF 116
 RESULT 10
 ID MDM2_XENLA STANDARD; PRT; 473 AA.
 AC P56273;

DT 15-JUL-1998 (rel. 36, Created)
 DT 30-MAY-2000 (rel. 39, Last sequence update)
 DT 01-MAR-2002 (rel. 41, Last annotation update)
 DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (P53-binding protein
 DE Mdm2) (Double minute 2 protein) (Xdm2).
 GN MDM2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97282706; PubMed-9136986;
 RA Marchal V., Elenbaas B., Taneyhill L., Piette J., Mechali M.,
 RA Nicolas J.-C., Levine A.J., Moreau J.;
 RT "Conservation of structural domains and biochemical activities of the
 RT MDM2 protein from *Xenopus laevis*.";
 RL Oncogene 14:1427-1433(1997).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 13-119 IN COMPLEX WITH P53.
 RX MEDLINE-97081050; PubMed-8675929;
 RA Kussie P.H., Gorina S., Marchal V., Elenbaas B., Moreau J.,
 RA Levine A.J., Pavletich N.P.;
 RT "Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor
 RT transactivation domain.";
 RL Science 274:948-953(1996).
 CC -1- FUNCTION: MAY BIND P53 PROTEIN AND MAY FUNCTION AS AN UBIQUITIN
 CC LIGASE E3.
 CC -1- COFACTOR: ZINC IS REQUIRED FOR UBIQUITIN LIGASE E3 ACTIVITY (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (by similarity).
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION INCREASES FROM OOCYTE STAGE I/II
 CC TO REACH ITS MAXIMUM IN OOCYTE STAGE V/VI IN UNFERTILIZED EGGS,
 CC AND THEN PROGRESSIVELY DECREASES TO BECOME UNDETECTABLE AT THE
 CC GASTRULA STAGE.
 CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 CC -1- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
 CC -1- SIMILARITY: BELONGS TO THE MDM2 / MDM4 FAMILY.
 DR PDB: 1YCO; 19-NOV-97.
 DR InterPro: IPR003160; MDM2.
 DR InterPro: IPR001876; Znf-RanBP.
 DR InterPro: IPR001841; Znf_ring.
 DR Pfam: PF02279; MDM2; 1.
 DR Pfam: PF00641; zf-RanBP; 1.
 DR SMART: SM00184; RING; 1.
 DR SMART: SM00547; ZNF_RBZ; 1.
 DR PROSITE: PS01358; ZF-RANBP2_1; 1.
 DR PROSITE: PS01099; ZF-RANBP2_2; 1.
 DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
 DR PROSITE: PS50089; ZF_RING_2; 1.
 KW Nuclear protein; Ligase; Ubiquitin conjugation; Zinc; Zinc-finger;
 KW Metal-binding; 3D-structure.
 FT DOMAIN 15 104 REGION I.
 FT DOMAIN 173 179 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 230 322 REGION II.
 FT DOMAIN 216 292 ASP/GLU-RICH (ACIDIC).
 FT ZN_FING 290 319 RANBP2-TYPE.
 FT ZN_FING 420 461 RING-TYPE.
 FT SEQUENCE 473 AA; 53464 MW; 7DA658DEB3BEE01 CRC64;
 Query Match 55.6%; Score 35; DB 1; Length 473;
 Best Local Similarity 61.5%; Pred. No. 52;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 KTEIISVYNXVAE 13
 Db 160 ETEIISVDHPAE 172
 RESULT 11
 ID MALO_STRPN STANDARD; PRT; 505 AA.
 ID MALO_STRPN

```

AC P29851;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase)
DE (disproportionating enzyme) (D-enzyme).
GN MALQ OR MALM OR SP2107.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=83129347; PubMed=6297760;
RA Lacks S.A., Dunn J.J., Greenberg B.;
RT "Identification of base mismatches recognized by the
RT heteroduplex-DNA-repair system of Streptococcus pneumoniae.";
RL Cell 31:327-336(1982).
[2]
RN SEQUENCE FROM N.A.
RX STRAIN=TIGR4;
RA Tetteijn H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Knout R., Molt A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Anguili S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.;
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser J.C.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
CC -1- CATALYTIC ACTIVITY: Transfers a segment of a (1,4)-alpha-D-glucan
CC to a new 4-position in an acceptor, which may be glucose or (1,4)-
CC alpha-D-glucan.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE DISPROPORTIONATING ENZYME FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: J01796; AAA26923.1; -
DR EMBL: AE007499; AAK76166.1; -
DR HSSP: 087172; 1CWY.
DR TIGR: SP2107; -
DR InterPro: IPR003385; 4A-glucanotrans.
DR Pfam: PF02446; 4A-glucanotrans.1.
KW Transferase; Glycosyltransferase; Carbohydrate metabolism;
KW Complete proteome.
SQ SEQUENCE 505 AA; 58076 MW; DA529A00A6D01A5 CRC64;

Query Match 55.6%; Score 35; DB 1; Length 505;
Best Local Similarity 54.5%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 EEISEVNXVAE 13
ID 1:1:1:1
DB 332 EELGELINTIAE 342

RESULT 12
PAC ARTVI STANDARD; PRT; 802 AA.
AC P31956;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Penicillin G acylase precursor (EC 3.5.1.11) (Penicillin G amidase)
DE (Penicillin G amidohydrolase).
GN PAC OR PA.
OS Arthrobacter viscosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=1673;
[1]
RN SEQUENCE FROM N.A.
RX STRAIN=ATCC 15294;
RX MEDLINE=94259306; PubMed=8200542;
RA Konstantinov M., Marjanovic N., Ljubicankic G., Glisin V.;
RT "The penicillin amidase of Arthrobacter viscosus (ATCC 15294).";
RL Gene 143:79-83(1994).
CC -1- CATALYTIC ACTIVITY: Penicillin + H(2)O = a fatty acid anion + 6-
CC antipenicillanate.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA CHAIN AND A BETA CHAIN
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Extracellular (potential).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S45 (SERINE PROTEASE) ALSO
CC KNOWN AS THE PENICILLIN ACYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: L04471; AAA22077.1; -
DR HSSP: P06875; 1AQ.
DR MEROPS: S45.001; -
DR InterPro: IPR002692; Penicill.amidase.
DR Pfam: PF01804; Penicill.amidase; 1.
KM Hydrolyase; Antibiotic resistance; zymogen; signal.
FT SIGNAL 1 24
FT CHAIN 25 802 POTENTIAL.
FT CHAIN 25 234 PENICILLIN G ACYLASE ALPHA SUBUNIT.
FT PROPEP 235 265 SPACER PEPTIDE.
FT CHAIN 266 802 PENICILLIN G ACYLASE BETA SUBUNIT.
FT ACT_SITE 266 266 BY SIMILARITY.
SQ SEQUENCE 802 AA; 92113 MW; 96635E8C526C85D CRC64;

Query Match 55.6%; Score 35; DB 1; Length 802;
Best Local Similarity 50.0%; Pred. No. 90;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 EEISEVNXVAE 14
ID 1:1:1:1
DB 548 EDINETNTYASF 559

RESULT 13
PAC_BACME STANDARD; PRT; 802 AA.
AC Q60136; Q9S463;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin G acylase precursor (EC 3.5.1.11) (Penicillin G amidase)
DE (Penicillin G amidohydrolase).
GN PAC OR PGA.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1404;
[1]
RN SEQUENCE FROM N.A.
RX STRAIN=ATCC 14945;
RA Kang J.H., Kim S.J., Park Y.C., Hwang Y., Yoo O.J., Kim Y.C.;
RT "Nucleotide sequence of the penicillin G acylase gene from
RT Bacillus megaterium and characteristics of the enzyme.";

```

```

RL Misalimurhag Holj 32:215-221(1994).
RP [2]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 14945;
RX MEDLINE-95180705; PubMed=7875576;
RA Martin L.M., Prieto A.M., Cortes E., Garcia J.L.;
RT "Cloning and sequencing of the pac gene encoding the penicillin G
RT acylase of Bacillus megaterium ATCC 14945."
RL FEMS Microbiol. Lett. 125:287-292(1995).
RN [3]
RP SEQUENCE FROM N.A.
RP STRAIN-CA4098;
RA Yang S., Huang X.D., Huang Y.H., Li S.Y., Yuan Z.Y.;
RT "Bacillus megaterium CA4098 penicillin G amidase structural gene
RT sequence."
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: Penicillin + H(2)O -> a fatty acid anion + 6-
CC ampenicillanate.
CC -1 SUBUNIT: HETERODIMER OF AN ALPHA CHAIN AND A BETA CHAIN
CC (BY SIMILARITY).
CC -1 SUBCELLULAR LOCATION: Extracellular (Potential).
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S45 (SERINE PROTEASE) ALSO
CC KNOWN AS THE PENICILLIN ACYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07682; AAB41343.1; -
DR EMBL: Z37542; CA85774.1; -
DR EMBL: AF161313; AAD45609.1; -
DR HSSP: P06875; IAO.
DR MEROPS: S45.001; -
DR InterPro: IPR002692; Penicill_amidase.
DR Pfam: PF01804; Penicill_amidase; 1.
KW Hydrolyase; Antibiotic resistance; zymogen; signal.
FT SIGNAL 1 24
FT CHAIN 25 802
FT CHAIN 25 234
FT PROPEP 235 265
FT CHAIN 266 802
FT ACT_SITE 266 266
FT VARIANT 3 3
FT VARIANT 224 224
FT VARIANT 232 232
FT VARIANT 254 254
FT VARIANT 349 349
FT VARIANT 470 470
FT VARIANT 524 524
FT VARIANT 569 569
FT VARIANT 586 586
FT VARIANT 657 657
FT VARIANT 740 740
FT VARIANT 789 791
FT SEQUENCE 802 AA; 91987 MW; 877CA0564E50DFD CRC64;
Query Match 55.6%; Score 35; DB 1; Length 802;
Best Local Similarity 50.0%; Pred. No. 90;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 3 EISEVNVXVAF 14
DB 548 EDINEINTASF 559
RCSA_KLEAF
ID RCSA_KLEAF STANDARD; PRT; 207 AA.
AC P05338;

```

```

DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Colanic acid capsular biosynthesis activation protein A.
GN RCSA.
OS Klebsiella aerogenes.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=28451;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88009858; PubMed=3309150;
RA Allen P., Hart C.A., Saunders J.R.;
RT "Isolation from Klebsiella and characterization of two rcs genes that
RT activate colanic acid capsular biosynthesis in Escherichia coli."
RL J. Gen. Microbiol. 133:331-340(1987).
CC -1 FUNCTION: POSITIVE REGULATOR OF CAPSULAR POLYSACCHARIDE
CC SYNTHESIS. RCSA AND RCSB FORM A COMPLEX TO PROMOTE TRANSCRIPTION
CC OF THE GENES FOR CAPSULE SYNTHESIS.
CC -1 SIMILARITY: BELONGS TO THE LUXR/DIRA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15748; AAA25142.1; -
DR PIR: A47615; A47615.
DR InterPro: IPR000792; HTH_LuxR.
DR Pfam: PF00196; Gete; 1.
DR PRINTS: PR00038; HTHLUXR.
DR SMART: SM00421; HTH_LUXR; 1.
DR PROSITE: PS00622; HTH_LUXR_FAMILY; 1.
KW Transcription regulation; DNA-binding; Activator.
FT DNA_BIND 155 174
FT SEQUENCE 207 AA; 23366 MW; 24E2253BB9DEE162 CRC64;
Query Match 54.0%; Score 34; DB 1; Length 207;
Best Local Similarity 53.8%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 KTEISEVNVXVAF 13
DB 26 KQELIVEVNSAD 38
RESULT 15
ID YK55_YEAST STANDARD; PRT; 307 AA.
AC P36155;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hypothetical 36.1 kDa protein in SIS2-MTD1 intergenic region.
GN YKR075C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95220693; PubMed=7705654;
RA di Como C.J., Rose R., Arndt K.T.;
RT "Overexpression of SIS2, which contains an extremely acidic region,
RT increases the expression of SWI4, CLN1 and CLN2 in sld mutants."
RL Genetics 139:95-107(1995).
RN [2]
RP SEQUENCE FROM N.A.
RP Pohl T.M., Pohl F.M.;

```

RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -! SIMILARITY: TO YEAST YOR062C.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U01878; AAA79998.1; -
DR EMBL; Z28300; CAA82154.1; -
DR PIR; S38152; S38152.
DR SCD; S0001783; YKR075C.
KW Hypothetical protein.
FT DOMAIN 254 270
SQ SEQUENCE 307 AA; 36113 MW; 0DC6D0241DADE1C5 CRC64;
HIS-RICH.
Query Match 54.0%; Score 34; DB 1; Length 307;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 TPEISEVNVAE 13
|||:|:|
DB 157 TEEVDEINASTE 168

Search completed: October 30, 2002, 12:27:47
Job time : 4.20393 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:30 ; Search time 4.36855 Seconds

(Without alignments)
554.401 Million cell updates/sec

Title: US-09-724-571-72

Perfect score: 63
Sequence: 1 KTEISEVKNVAEF 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOIST:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	76.2	79	11	O35463
2	48	76.2	82	4	P78438
3	48	76.2	82	4	O16014
4	48	76.2	82	4	O16019
5	48	76.2	82	4	O16020
6	48	76.2	607	11	O99K32
7	48	76.2	695	6	O95KN7
8	48	76.2	695	11	P97487
9	48	76.2	695	11	O60496
10	48	76.2	770	6	O9TU10
11	47	74.6	534	13	O93296
12	47	74.6	569	13	O9PVL1
13	47	74.6	695	13	O9DCJ8
14	47	74.6	751	13	O9DGJ7
15	42	66.7	693	13	O98SG0
16	42	66.7	747	13	O91963

17	39	61.9	699	13	O57394	O57394 narke japon
18	38	60.3	316	5	O9XVK7	O9XVK7 caenorhabdi
19	38	60.3	335	5	O96847	O96847 dictyosteli
20	38	60.3	335	5	O23896	O23896 dictyosteli
21	38	60.3	1354	5	O9V621	O9V621 drosophila
22	38	60.3	1354	5	O9V6K1	O9V6K1 drosophila
23	38	60.3	1803	5	O9VU16	O9VU16 drosophila
24	37.5	59.5	213	9	O9XJ88	O9XJ88 bacterioph
25	37	58.7	561	5	O950M4	O950M4 caenorhabdi
26	37	58.7	564	9	O9AZ64	O9AZ64 lactococcus
27	37	58.7	568	4	O96LQ3	O96LQ3 homo sapien
28	37	58.7	946	10	O9FF91	O9FF91 arabidopsis
29	36.5	57.9	587	10	O9LYU1	O9LYU1 arabidopsis
30	36	57.1	180	16	O92P99	O92P99 rhizobium m
31	36	57.1	190	12	O9EMT2	O9EMT2 amsaecta moo
32	36	57.1	354	5	O9VER2	O9VER2 drosophila
33	36	57.1	415	5	O44228	O44228 halocynthia
34	36	57.1	470	8	O9MSU0	O9MSU0 juniperus c
35	36	57.1	502	8	O9MST8	O9MST8 juniperus p
36	36	57.1	512	10	O9C8H8	O9C8H8 arabidopsis
37	36	57.1	513	10	O38821	O38821 arabidopsis
38	36	57.1	570	10	O9PCT6	O9PCT6 caenorhabdi
39	36	57.1	650	5	O17866	O17866 caenorhabdi
40	36	57.1	695	13	O98SE9	O98SE9 xenopus lae
41	36	57.1	778	5	O9U9K6	O9U9K6 caenorhabdi
42	36	57.1	837	5	O25751	O25751 plasmodium
43	36	57.1	864	5	O93336	O93336 caenorhabdi
44	36	57.1	867	10	O81747	O81747 arabidopsis
45	36	57.1	1117	5	O9U9K7	O9U9K7 caenorhabdi

ALIGNMENTS

Result 1
ID O35463 PRELIMINARY; PRT; 79 AA.
AC O35463;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ALHEIMER'S AMYLOID BETA PROTEIN (FRAGMENT).
GN BETA APP.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Sambamurti K., Plinix I., Gandhi S.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF030413; AAB86608.1; -.
DR HSSP: P05067; IBA4.
FT NON_TER
FT NON_TER
SQ SEQUENCE 79 AA: 8538 MW: 37F2C6C3BFF3F597 CRC64;
Query Match 76.2%; Score 48; DB 11; Length 79;
Best Local Similarity 78.6%; Pred. NO. 0.19;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 KTEISEVKNVAEF 14
Db 11 KTEISEVKNVAEF 24
RESULT 2
ID P78438 PRELIMINARY; PRT; 82 AA.
AC P78438;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)

```

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE AMYLOID PROTEIN (BETA-AMYLOID PROTEIN) (FRAGMENT).
GN APP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89392030; PubMed=2675837;
RA Johnstone E.M., Chaney M.O., Moore R.E., Ward K.E., Norris F.H.,
RA Little S.P.'s disease amyloid peptide is encoded by two exons and shows
RT similarity to soybean trypsin inhibitor.";
RL Biochem. Biophys. Res. Commun. 163:1248-1255(1989).
RN [2]
RP SEQUENCE OF 19-48 FROM N.A.
RX MEDLINE=87120329; PubMed=2949367;
RA Tanzil R.E., Gusella J.F., Watkins P.C., Bruns G.A., George-Hyslop P.,
RA Van Keuren M.L., Patterson D., Pagan S., Kuralt D.M., Neve R.L.;
RT "Amyloid beta protein gene: cDNA, mRNA distribution, and genetic
RT linkage near the Alzheimer locus.";
RL Science 235:880-884(1987).
RN [3]
RP SEQUENCE OF 32-63 FROM N.A.
RX MEDLINE=93035397; PubMed=1415269;
RA Kamino K., Orr H.T., Payant H., Wjisman E.M., Alonso M.E., Pulst S.M.,
RA Anderson L., O'dahl S., Nemens E., White J.A.;
RT "Linkage and mutational analysis of familial Alzheimer disease
RT kindreds for the APP gene region.";
RL Am. J. Hum. Genet. 51:998-1014(1992).
DR EMBL: M29270; AAA51768.1; -.
DR EMBL: M29269; AAA51768.1; JOINED.
DR EMBL: M15532; AAA51564.1; -.
DR EMBL: S45136; AAB23646.1; -.
DR HSSP: P05067; 1BA4.
FT NON_TER 1
SQ SEQUENCE 82 AA; 8994 MW; 8DA9E42B813A070E CRC64;

Query Match 76.2%; Score 48; DB 4; Length 82;
Best Local Similarity 78.6%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
DB 7 KTEISEVKNDAEF 20
ID 016014 PRELIMINARY; PRT; 82 AA.
AC 016014;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BETA-AMYLOID PEPTIDE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93236601; PubMed=8476439;
RA Denman R.B., Rosenzwaig R., Miller D.L.;
RT "A system for studying the effect(s) of familial Alzheimer disease
RT mutations on the processing of the beta-amyloid peptide precursor.";
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
DR EMBL: S60721; AAB26263.2; -.
DR HSSP: P05067; 1BA4.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8972 MW; F534AA5B3EA9230A CRC64;

```

```

Query Match 76.2%; Score 48; DB 4; Length 82;
Best Local Similarity 78.6%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
DB 8 KTEISEVKNDAEF 21
ID 016019 PRELIMINARY; PRT; 82 AA.
AC 016019;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BETA-AMYLOID PEPTIDE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93236601; PubMed=8476439;
RA Denman R.B., Rosenzwaig R., Miller D.L.;
RT "A system for studying the effect(s) of familial Alzheimer disease
RT mutations on the processing of the beta-amyloid peptide precursor.";
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
DR EMBL: S61380; AAB26264.2; -.
DR HSSP: P05067; 1BA4.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8938 MW; F534AA50E579230A CRC64;

Query Match 76.2%; Score 48; DB 4; Length 82;
Best Local Similarity 78.6%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
DB 8 KTEISEVKNDAEF 21
ID 016020 PRELIMINARY; PRT; 82 AA.
AC 016020;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BETA-AMYLOID PEPTIDE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93236601; PubMed=8476439;
RA Denman R.B., Rosenzwaig R., Miller D.L.;
RT "A system for studying the effect(s) of familial Alzheimer disease
RT mutations on the processing of the beta-amyloid peptide precursor.";
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
DR EMBL: S61383; AAB26265.2; -.
DR HSSP: P05067; 1BA4.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8882 MW; F534AA5AE5D9230A CRC64;

Query Match 76.2%; Score 48; DB 4; Length 82;
Best Local Similarity 78.6%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```


Oy 1 KTEISEVNXVAEF 14
 ||||||| |||
 Db 8 KTEISEVKMDAEF 21

RESULT 6

099K32 ID 099K32 PRELIMINARY; PRT; 607 AA.

AC 099K32, 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 OS HYPOTHETICAL 68.4 KDA PROTEIN (FRAGMENT).

OC Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;

RC SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY TUMOR. WAP-TGF ALPHA MODEL. 7 MONTHS OLD, GROSS
 RC TISSUE.

RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC005490; AA05490.1; -
 DR HSSP: P05067; IAPP.
 DR InterPro: IPR001868; A4_APP.
 DR Pfam: PF00014; Kunitz_BPTI.
 DR PRINTS: PR00203; AMYLOIDA4.
 DR PRINTS: PR00759; BASICPTASE.
 DR SMART: SM00131; KU; 1.

DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE: PS0279; BPTI_KUNITZ_2; 1.
 KW Hypothetical protein; Serine protease inhibitor.
 FT NON-TER 1

SQ SEQUENCE 607 AA; 68391 MW; BF02214CBA7D172 CRC64;

Query Match

Best Local Similarity 76.2%; Score 48; DB 11; Length 607;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KTEISEVNXVAEF 14
 ||||||| |||
 Db 499 KTEISEVKMDAEF 512

RESULT 7

095KN7 ID 095KN7 PRELIMINARY; PRT; 695 AA.

AC 095KN7, 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 OS AMYLOID B-PROTEIN PRECURSOR.

OC Macaca fascicularis (Cebus eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;

RP SEQUENCE FROM N.A.
 RC TISSUE=CEREBELLUM;
 RX MEDLINE=91273117; PubMed=1905108;

RA Podlisy M.B., Tolian D.R., Selkoe D.J.;
 RT "Homology of the amyloid beta protein precursor in monkey and human
 RT supports a primate model for beta amyloidosis in Alzheimer's
 RT disease";

RL Am. J. Pathol. 138:1423-1435(1991).
 DR EMBL: M58727; AAA36829.1; -
 FT SIGNAL 1 17 POTENTIAL.

FT CHAIN 597 636 POTENTIAL.
 SQ SEQUENCE 695 AA; 78663 MW; 4FE6A0139F69D56 CRC64;

Query Match

Best Local Similarity 76.2%; Score 48; DB 6; Length 695;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KTEISEVNXVAEF 14
 ||||||| |||
 Db 587 KTEISEVKMDAEF 600

RESULT 8

097487 ID 097487 PRELIMINARY; PRT; 695 AA.

AC 097487, 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 OS HIPPOCAMPAL AMYLOID PROTEIN.

OC Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RC SEQUENCE FROM N.A.
 RC STRAIN=SAMP8; TISSUE=HIPPOCAMPUS;
 RA Flood J.F., Kumar V.B., Sasser T., Word I., Morley J.E.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE OF 581-662 FROM N.A.
 RC STRAIN=129SV;
 RA Wragg M.A., Busfield F., Duff K., Korenblat K., Capecchi M.,
 RL Loring J.F., Goate A.M.;
 DR Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U84012; AAB41502.1; -
 DR EMBL: U82624; AAB40919.1; -
 DR HSSP: P05067; IMWP.
 DR MGD: MGI:88059; APP.

DR InterPro: IPR001868; A4_APP.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR PRINTS: PR00203; AMYLOIDA4.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.

SQ SEQUENCE 695 AA; 78414 MW; 9A5FB2ED261236E CRC64;

Query Match

Best Local Similarity 76.2%; Score 48; DB 11; Length 695;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 KTEISEVNXVAEF 14
 ||||||| |||
 Db 587 KTEISEVKMDAEF 600

RESULT 9

060496 ID 060496 PRELIMINARY; PRT; 695 AA.

AC 060496, 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 OS PUTATIVE AMYLOID PRECURSOR PROTEIN.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.
 OX NCBI_TaxID=10143;

RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=97236426; PubMed=9116031;

RA Beck M., Mueller D., Bigl V.
 RT "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
 alternative splicing."
 RL Biochim. Biophys. Acta 1351:17-21(1997).
 DR EMBL: X97631; CAA66230.1; -.
 DR HSSP: P05067; 1BA4.
 DR InterPro: IPR001868; A4_APP.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR PRINTS: PR00203; AMYLOIDA4.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR PROSITE: PS00319; A4_INTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 SO SEQUENCE 695 AA; 78701 MW; 5196A0C4017F16AB CRC64;

Query Match 76.2%; Score 48; DB 11; Length 695;
 Best Local Similarity 78.6%; Pred. No. 1.7;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNXVAEF 14
 |||||
 DB 587 KTEISEVXKMAEF 600

RESULT 10
 O9TUI0 PRELIMINARY; PRT; 770 AA.
 AC O9TUI0:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE AMYLOID PRECURSOR PROTEIN.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kimura A., Takahashi T.
 RT "Amyloid Precursor Protein 770."
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB032550; BAA84580.1; -.
 DR HSSP: P05067; 1AAP.
 DR InterPro: IPR001868; A4_APP.
 DR InterPro: IPR002223; Kunitz_BPTI.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR Pfam: PF00014; Kunitz_BPTI; 1.
 DR PRINTS: PR00203; AMYLOIDA4.
 DR PRINTS: PR00759; BASICPTASE.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR SMART: SM00131; K0; 1.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
 KM Serine protease inhibitor.
 SO SEQUENCE 770 AA; 86961 MW; 5F7A1DCB2BC583E CRC64;

Query Match 76.2%; Score 48; DB 6; Length 770;
 Best Local Similarity 78.6%; Pred. No. 1.8;
 Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNXVAEF 14
 |||||
 DB 662 KTEISEVXKMAEF 675

RESULT 11
 O93296 PRELIMINARY; PRT; 534 AA.
 AC O93296:
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE AMYLOID PROTEIN (FRAGMENT).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98337885; PubMed=9671674;
 RA Barnes N.Y., Li L., Yoshikawa K., Schwartz L.M., Oppenheim R.W.,
 RA Milligan C.E.;
 RT "Increased production of amyloid precursor protein provides a
 substrate for caspase-3 in dying motoneurons."
 RL J. Neurosci. 18:5869-5880(1998).
 DR EMBL: AF042098; AAC25052.1; -.
 DR HSSP: P05067; 1BA4.
 DR InterPro: IPR001868; A4_APP.
 DR PRINTS: PR00203; AMYLOIDA4.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 FT NON TER 1
 SO SEQUENCE 534 AA; 60597 MW; FB53ECC2E66D4C92 CRC64;

Query Match 74.6%; Score 47; DB 13; Length 534;
 Best Local Similarity 71.4%; Pred. No. 2;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNXVAEF 14
 |||||
 DB 426 KTEISEVXKMAEF 439

RESULT 12
 O9PVL1 PRELIMINARY; PRT; 569 AA.
 AC O9PVL1:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE AMYLOID PROTEIN (FRAGMENT).
 GN APP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Coulson E.J., Beyreuther K., Masters C.L.;
 RT "What the evolution of the amyloid protein precursor supergene family
 tells us about its function."
 RL Neurochem. Int. 0:0-0(2000).
 DR EMBL: AF030341; AAF12698.1; -.
 DR HSSP: P05067; 1BA4.
 DR InterPro: IPR001868; A4_APP.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR PRINTS: PR00203; AMYLOIDA4.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 FT NON TER 1
 SO SEQUENCE 569 AA; 64753 MW; 0AB8BB851863A19D CRC64;

Query Match 74.6%; Score 47; DB 13; Length 569;
 Best Local Similarity 71.4%; Pred. No. 2.1;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KTEISEVNXVAEF 14
 |||||
 DB 462 KTEISEVXKMAEF 475

RESULT 13

Q9DCJ8 PRELIMINARY: PRT: 695 AA.
 AC Q9DCJ8;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BETA-AMYLOID PRECURSOR PROTEIN 695 ISOFORM.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sarasa M., Rodolosse A., Sorribas V.;
 RT "Cloning of full-length chicken beta-amyloid precursor protein
 isoforms."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF289218; AAG00593.1; -;
 DR HSSP: P05067; 1BA4.
 DR InterPro: IPR001868; A4_APP.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR PRINTS: PRO0203; AMYLOIDA4.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR SEQUENCE 695 AA; 78565 MW; F201ED02AEC86D95 CRC64;

Query Match

Best Local Similarity 74.6%; Score 47; DB 13; Length 695;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
 DB 587 KTEEVEVKMDAEF 600

RESULT 14

Q9DCJ7 PRELIMINARY: PRT: 751 AA.
 AC Q9DCJ7;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BETA-AMYLOID PRECURSOR PROTEIN 751 ISOFORM.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sarasa M., Rodolosse A., Sorribas V.;
 RT "Cloning of full-length chicken beta-amyloid precursor protein
 isoforms."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF289219; AAG00594.1; -;
 DR HSSP: P05067; 1BA4.
 DR InterPro: IPR001868; A4_APP.
 DR InterPro: IPR002223; Kunitz_BPTI.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR PRINTS: PRO0203; AMYLOIDA4.
 DR PRINTS: PRO0759; BASICTPASE.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR SMART: SM00131; KU; 1.
 DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE: PS50279; BPTI_KUNITZ_2; 1.
 DR Serine protease inhibitor.
 DR SEQUENCE 751 AA; 84705 MW; E78E9413A8033D84 CRC64;

Query Match 74.6%; Score 47; DB 13; Length 751;
 Best Local Similarity 71.4%; Pred. No. 2.8;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
 DB 643 KTEEVEVKMDAEF 656

RESULT 15

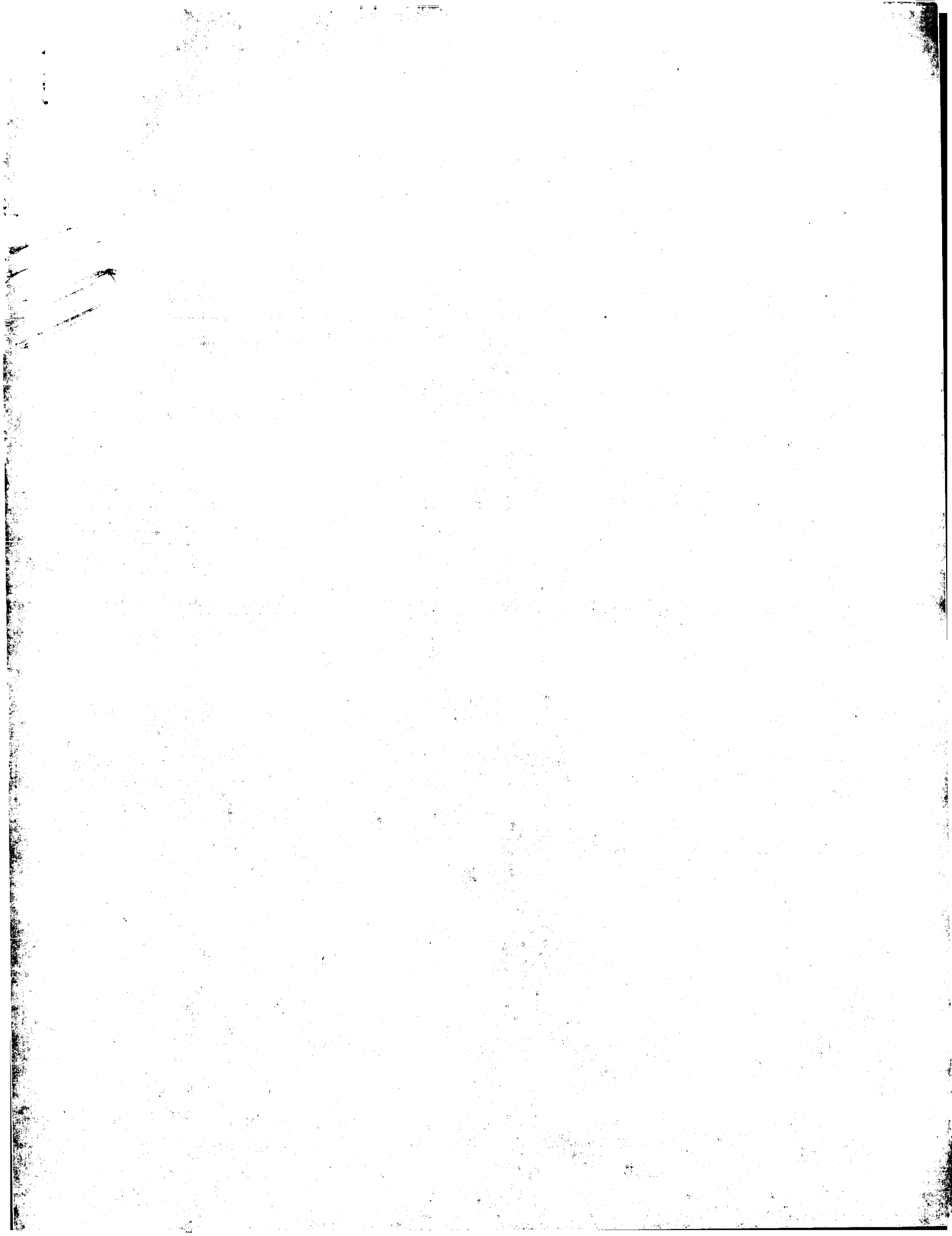
Q98SG0 PRELIMINARY: PRT: 693 AA.
 AC Q98SG0;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BETA-AMYLOID PRECURSOR PROTEIN A.
 GN APP.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8395;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Van den Hurk W.H.;
 RL Thesis (2001), Department of Biological Sciences,
 RL University of Nijmegen, Nijmegen, Netherlands.
 DR EMBL: AJ288150; CAC37193.1; -;
 DR HSSP: P05067; 1H23.
 DR InterPro: IPR001868; A4_APP.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR PRINTS: PRO0203; AMYLOIDA4.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 DR Signal.
 DR FT SIGNAL.
 DR SEQUENCE 693 AA; 78568 MW; CAF1DF655C1AB653 CRC64;

Query Match

Best Local Similarity 66.7%; Score 42; DB 13; Length 693;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KTEISEVNXVAEF 14
 DB 585 KTEEVEVKMDSEY 598

Search completed: October 30, 2002, 12:30:01
 Job time : 6.36855 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:24:20 ; Search time 0.997543 Seconds
(without alignments)
171,400 Million cell updates/sec

Title: US-09-724-571-78

Perfect score: 29

Sequence: 1 VMXAEF 7

Scoring table:

BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCtus.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	82.8	972	3 US-08-335-844A-24	Sequence 24, Appl
2	23	79.3	54	2 US-08-456-647B-18	Sequence 18, Appl
3	23	79.3	54	2 US-08-456-647B-46	Sequence 46, Appl
4	23	79.3	54	2 US-08-456-647B-47	Sequence 47, Appl
5	23	79.3	54	2 US-08-237-401A-18	Sequence 18, Appl
6	23	79.3	54	2 US-08-237-401A-46	Sequence 46, Appl
7	23	79.3	54	2 US-08-237-401A-47	Sequence 47, Appl
8	23	79.3	299	2 US-08-701-191A-13	Sequence 13, Appl
9	23	79.3	300	2 US-08-701-191A-31	Sequence 31, Appl
10	23	79.3	309	2 US-08-701-191A-9	Sequence 9, Appl
11	23	79.3	310	2 US-08-701-191A-1	Sequence 1, Appl
12	23	79.3	310	2 US-08-701-191A-6	Sequence 6, Appl
13	23	79.3	310	2 US-08-701-191A-7	Sequence 7, Appl
14	23	79.3	310	2 US-08-701-191A-8	Sequence 8, Appl
15	23	79.3	313	1 US-08-278-089A-17	Sequence 17, Appl
16	23	79.3	313	1 US-08-838-957A-16	Sequence 16, Appl
17	23	79.3	315	2 US-08-701-191A-2	Sequence 2, Appl
18	23	79.3	318	2 US-08-701-191A-10	Sequence 10, Appl
19	23	79.3	351	2 US-08-701-191A-3	Sequence 3, Appl
20	23	79.3	378	1 US-08-070-165F-8	Sequence 8, Appl
21	23	79.3	378	1 US-08-885-418-8	Sequence 8, Appl
22	23	79.3	388	1 US-08-070-165F-4	Sequence 4, Appl
23	23	79.3	388	1 US-08-885-418-4	Sequence 4, Appl
24	23	79.3	612	2 US-08-673-789-11	Sequence 11, Appl
25	23	79.3	652	1 US-08-471-570-10	Sequence 10, Appl
26	23	79.3	729	1 US-07-640-029-3	Sequence 3, Appl
27	23	79.3			

28	23	79.3	729	1 US-08-070-165F-6	Sequence 6, Appl
29	23	79.3	729	2 US-08-885-418-6	Sequence 6, Appl
30	23	79.3	731	1 US-07-921-807B-5	Sequence 5, Appl
31	23	79.3	731	1 US-08-070-165F-10	Sequence 10, Appl
32	23	79.3	731	1 US-08-441-944A-5	Sequence 5, Appl
33	23	79.3	731	2 US-08-885-418-10	Sequence 10, Appl
34	23	79.3	733	1 US-08-439-992A-3	Sequence 3, Appl
35	23	79.3	733	1 US-07-640-029-4	Sequence 4, Appl
36	23	79.3	733	1 US-07-921-807B-6	Sequence 6, Appl
37	23	79.3	733	1 US-08-441-944A-6	Sequence 6, Appl
38	23	79.3	733	4 US-08-439-992A-4	Sequence 4, Appl
39	23	79.3	769	1 US-08-471-570-8	Sequence 8, Appl
40	23	79.3	801	4 US-09-383-630-6	Sequence 6, Appl
41	23	79.3	802	4 US-09-173-151A-33	Sequence 3, Appl
42	23	79.3	806	4 US-09-383-630-3	Sequence 3, Appl
43	23	79.3	816	1 US-07-640-029-1	Sequence 1, Appl
44	23	79.3	817	1 US-07-640-029-2	Sequence 2, Appl
45	23	79.3	820	1 US-07-921-807B-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-335-844A-24
Sequence 24, Application US/08335844A
Patent No. 6066503
GENERAL INFORMATION:
APPLICANT: GRAHAM, MARGARET
APPLICANT: SMITH, TREVOR STANLEY
APPLICANT: MUNN, EDWARD ALBERT
APPLICANT: KNOX, DAVID PATRICK
APPLICANT: OLIVER, JOANNA JANE
APPLICANT: NEWTON, SUSAN ELIZABETH
TITLE OF INVENTION: RECOMBINANT DNA MOLECULES ENCODING
TITLE OF INVENTION: AMINOPEPTIDASE ENZYMES AND THEIR USE
TITLE OF INVENTION: VACCINES AGAINST HELMINTH INFECTIONS
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rothwell, Figg, Ernst & Kurz
STREET: Suite 701-E, 555 Thirteenth St., N.W
CITY: Washington
STATE: D. C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/335,844A
FILING DATE: 09-JAN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB PCT/GB93/00943
FILING DATE: 06-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9209936
FILING DATE: 08-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: WALKER, Barbara W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 1181-223A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 972 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-335-844A-24

Query Match 82.8%; Score 24; DB 3; Length 972;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7

Db 542 VISVAEF 548

RESULT 2

US-08-456-647B-18

; Sequence 18, Application US/08456647B

; Patent No. 5811516

; GENERAL INFORMATION:

; APPLICANT: Lemke Ph.D. et al., Greg E.

; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/456,647B

; FILING DATE: 02-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/237,401

; FILING DATE: 02-MAY-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/884,486

; FILING DATE: 15-MAY-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Wetherell Ph.D., John R.

; REGISTRATION NUMBER: 31,678

; REFERENCE/DOCKET NUMBER: 07251/007002

; TELEPHONE: (619) 678-5070

; TELEFAX: (619) 678-5099

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 54 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-456-647B-18

Query Match 79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 28;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7

Db 9 VMKIADF 15

RESULT 3

US-08-456-647B-46

; Sequence 46, Application US/08456647B

; Patent No. 5811516

; GENERAL INFORMATION:

; APPLICANT: Lemke Ph.D. et al., Greg E.

; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES

; NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,647B

FILING DATE: 02-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/237,401

FILING DATE: 02-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/884,486

FILING DATE: 15-MAY-1992

ATTORNEY/AGENT INFORMATION:

NAME: Wetherell Ph.D., John R.

REGISTRATION NUMBER: 31,678

REFERENCE/DOCKET NUMBER: 07251/007002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 678-5070

TELEFAX: (619) 678-5099

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 54 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match 79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 28;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7

Db 9 VMKIADF 15

RESULT 4

US-08-456-647B-47

; Sequence 47, Application US/08456647B

; Patent No. 5811516

; GENERAL INFORMATION:

; APPLICANT: Lemke Ph.D. et al., Greg E.

; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/456,647B

; FILING DATE: 02-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/237,401

;; FILING DATE: 02-MAY-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/884,486
;; FILING DATE: 15-MAY-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Wetterell Ph.D., John R.
;; REGISTRATION NUMBER: 31,678
;; REFERENCE/DOCKET NUMBER: 07251/007002
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619) 678-5070
;; TELEFAX: (619) 678-5099
;; INFORMATION FOR SEQ ID NO: 47:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 54 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-456-647B-47

Query Match 79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 28;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||:|:|
DB 9 VKMIADF 15

RESULT 5
US-08-237-401A-18
; Sequence 18, Application US/08237401A
; Patent No. 5837448
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,401A
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-237-401A-18

Query Match 79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 28;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
||:|:|
DB 9 VKMIADF 15

RESULT 6
US-08-237-401A-46
; Sequence 46, Application US/08237401A
; Patent No. 5837448
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,401A
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-237-401A-46

Query Match 79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 28;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
||:|:|
DB 9 VKMIADF 15

RESULT 7
US-08-237-401A-47
; Sequence 47, Application US/08237401A
; Patent No. 5837448
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US

ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/237,401A
FILING DATE: 02-May-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-May-1992
ATTORNEY/AGENT INFORMATION:
NAME: Halle Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-237-401A-47

Query Match 79.3%; Score 23; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 28;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXYAEF 7
||:|:|
Db 9 VMKIADF 15

RESULT 8
US-08-701-191A-13
Sequence 13, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
and Steven R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088

TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 299 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-13

Query Match 79.3%; Score 23; DB 2; Length 299;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXYAEF 7
||:|:|
Db 172 VMKIADF 178

RESULT 9
US-08-701-191A-31
Sequence 31, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
and Steven R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 955-0440
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-31

Query Match 79.3%; Score 23; DB 2; Length 300;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
11 :1:1
Db 172 VMK1ADF 187

RESULT 10

US-08-701-191A-9
Sequence 9, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
and Stevan R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-9

Query Match 79.3% Score 23; DB 2; Length 309;
Best Local Similarity 57.1% Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
11 :1:1
Db 181 VMK1ADF 187

RESULT 11
US-08-701-191A-1
Sequence 1, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
and Stevan R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-701-191A-1

Query Match 79.3% Score 23; DB 2; Length 310;
Best Local Similarity 57.1% Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
11 :1:1
Db 181 VMK1ADF 187

RESULT 12
US-08-701-191A-6
Sequence 6, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
and Stevan R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-6

Query Match
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
||:|:|
Db 181 VMKIADF 187

RESULT 13
US-08-701-191A-7
Sequence 7, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
APPLICANT: and Steven R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-7

Query Match
Best Local Similarity 79.3%; Score 23; DB 2; Length 310;
Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
||:|:|
Db 181 VMKIADF 187

RESULT 14
US-08-701-191A-8
Sequence 8, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
APPLICANT: and Steven R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-8

Query Match
Best Local Similarity 79.3%; Score 23; DB 2; Length 310;
Pred. No. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
||:|:|
Db 181 VMKIADF 187

```

RESULT 15
US-09-390-326-7
; Sequence 7, Application US/09390326
; Patent No. 6316603
; GENERAL INFORMATION:
; APPLICANT: MCTIGUE, MICHELE A.
; APPLICANT: WICKERSHAM, JOHN A.
; APPLICANT: PINKO, CHRIS
; APPLICANT: SHOWALTER, RICHARD
; APPLICANT: PARAST, CAMRAN V.
; APPLICANT: TEMPCZYK-RUSSELL, ANNA
; APPLICANT: GEHRING, MICHAEL R.
; APPLICANT: MROCKOWSKI, BARBARA
; APPLICANT: KAN, CHEN-CHEN
; APPLICANT: VILAFRANCA, J. ERNEST
; APPLICANT: APPELT, KRYSZTOF
; TITLE OF INVENTION: MODIFICATIONS OF THE VEGF RECEPTOR-2 PROTEIN AND
; FILE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 0125-0016US
; CURRENT APPLICATION NUMBER: US/09/390,326
; CURRENT FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-390-326-7

```

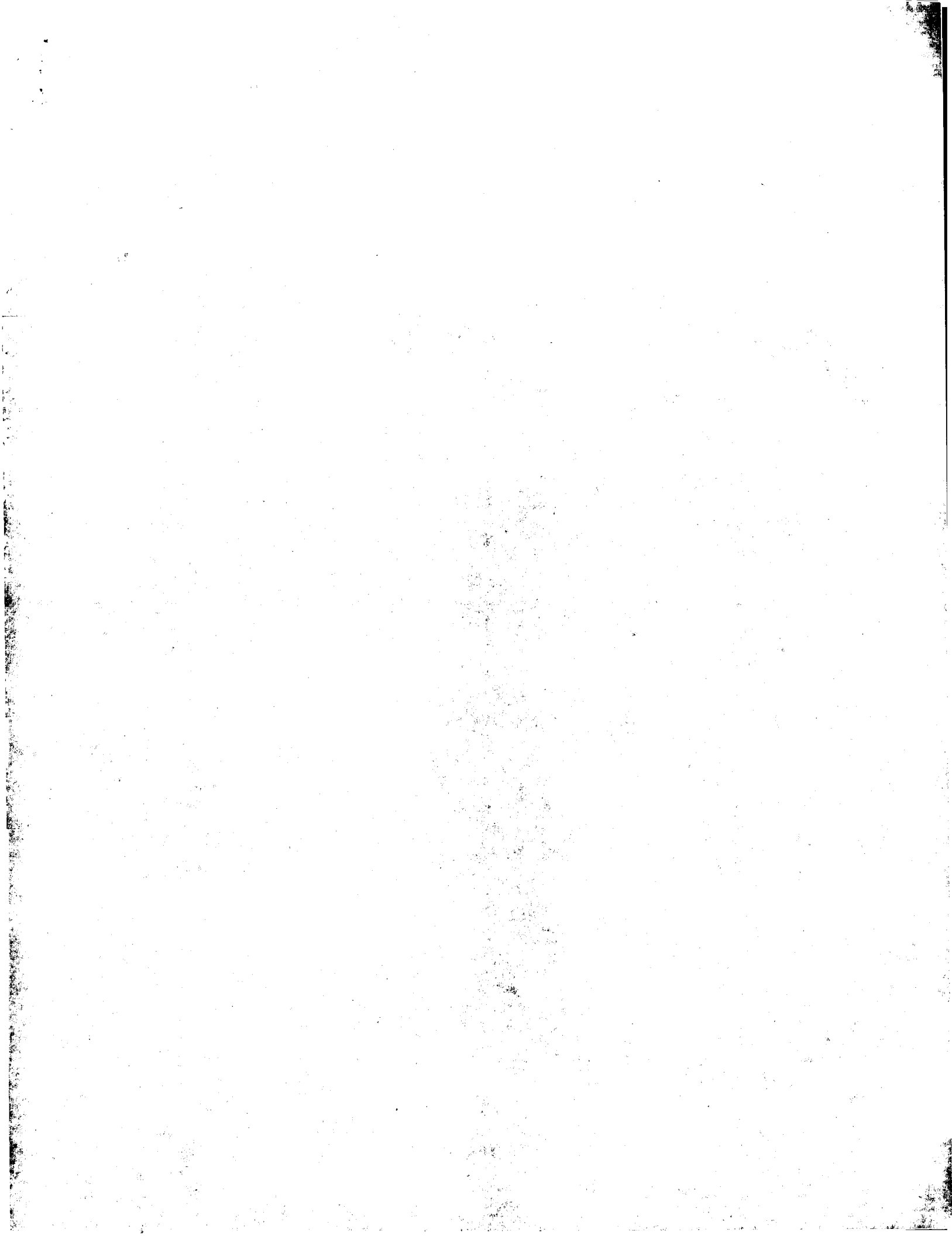
```

Query Match          79.3% Score 23; DB 4; Length 310;
Best Local Similarity 57.1%; Pred. NO. 1.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VMXVAEF 7
      ||:|:|
Db      181 VMKIADF 187

```

Search completed: October 30, 2002, 12:32:33
Job time : 1.99754 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:20:50 ; Search time 2.80344 Seconds

(without alignments)
277.344 Million cell updates/sec

Title: US-09-724-571-78

Perfect score: 29

Sequence: 1 VMXVAEF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_032802:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	27	93.1	7	21	AA807871	A beta-secretase 1
2	27	93.1	8	21	AA807872	A beta-secretase 1
3	26	89.7	821	20	AAV34479	Porphyromonas ging
4	26	89.7	869	20	AAV34354	Porphyromonas ging
5	24	82.8	75	22	AAU56778	Propionibacterium
6	24	82.8	132	22	AA41984	Human polypeptide
7	24	82.8	132	22	AA41985	Human polypeptide
8	24	82.8	132	22	AA41986	Human polypeptide
9	24	82.8	225	19	AAV85820	S. pneumoniae deri
10	24	82.8	236	22	AAU50802	Propionibacterium
11	24	82.8	254	20	AAV36930	Amino acid sequenc

12	24	82.8	286	22	ABG18334	Novel human diageno
13	24	82.8	335	22	AAU37753	Streptococcus pneu
14	24	82.8	361	21	AA820945	Arabidopsis thalia
15	24	82.8	361	21	AA824458	Arabidopsis thalia
16	24	82.8	361	21	AA845883	Arabidopsis thalia
17	24	82.8	373	22	AA440200	Human polypeptide
18	24	82.8	378	21	AA820944	Arabidopsis thalia
19	24	82.8	378	21	AA824457	Arabidopsis thalia
20	24	82.8	378	21	AA845882	Arabidopsis thalia
21	24	82.8	384	21	AA71036	Human membrane tra
22	24	82.8	384	22	AA40198	Human polypeptide
23	24	82.8	384	22	AA40199	Human polypeptide
24	24	82.8	414	21	AA820943	Arabidopsis thalia
25	24	82.8	414	21	AA824456	Arabidopsis thalia
26	24	82.8	414	21	AA845881	Arabidopsis thalia
27	24	82.8	584	22	AA862295	Putative P. abyssal
28	24	82.8	817	22	AB870027	Drosophila melanog
29	24	82.8	962	15	AA858701	Aminopeptidase H11
30	24	82.8	972	14	AA851280	Helminth aminopept
31	23	79.3	36	18	AAW27782	UDP-N-acetylglucos
32	23	79.3	54	19	AAW79151	Receptor protein t
33	23	79.3	54	20	AAW81408	Receptor protein t
34	23	79.3	84	22	AAU41413	Propionibacterium
35	23	79.3	90	22	AA887752	Human T2R22 amino
36	23	79.3	143	21	AA810816	Archaeoglobus fulg
37	23	79.3	150	22	AAU33403	Enterococcus faeca
38	23	79.3	189	22	AA852817	Escherichia coli p
39	23	79.3	237	22	AA890704	C glutamicum prote
40	23	79.3	246	22	AA858040	Drosophila melanog
41	23	79.3	249	22	AAE13127	Sheep SLS FGFR3 no
42	23	79.3	249	22	AAE13136	Sheep SLS FGFR3 mu
43	23	79.3	280	22	AAE30751	C glutamicum prote
44	23	79.3	287	20	AAV03631	Amino acid sequenc
45	23	79.3	299	22	AB863276	Drosophila melanog

ALIGNMENTS

RESULT 1	AA807871	standard; peptide: 7 AA.
XX	AA807871:	
AC	14-NOV-2000	(first entry)
XX		
DT		
XX		
DE		A beta-secretase inhibitor peptide.
XX		
KW		Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW		amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW		inhibitor.
XX		
OS		Synthetic.
XX		
FM		
FT	Key	Location/Qualifiers
XX	Modified-site	3 /note="hydroxyethylene"
XX		
PN	W0200047618-A2.	
XX		
PD	17-AUG-2000.	
XX		
PF	10-FEB-2000; 2000WO-US03819.	
XX		
PR	10-FEB-1999; 99US-0119571.	
XX	15-JUN-1999; 99US-0139172.	
XX		
PA	(ELAN-) ELAN PHARM INC.	
XX		
XX	Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;	
PI	Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;	
XX		

DR WPI; 2000-533011/48.
XX Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
PS Disclosure; Page 12; 121pp; English.
XX The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents an inhibitor of beta-secretase enzyme.
XX
SQ Sequence 7 AA;
Query Match 93.1%; Score 27; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VMXVAEF 7
DB 1 VMXVAEF 7
RESULT 2
AAB07872
ID AAB07872 standard; peptide; 8 AA.
XX AAB07872;
XX 14-NOV-2000 (first entry)
XX A beta-secretase inhibitor peptide.
XX
XX Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 4
FT /note="hydroxyethylene"
FT
XX
XX WO200047618-A2.
XX
XX 17-AUG-2000.
XX
XX 10-FEB-2000; 2000WO-US03819.
XX
XX 10-FEB-1999; 99US-0119571.
XX
XX 15-JUN-1999; 99US-0139172.
XX
XX (ELAN-) ELAN PHARM INC.
XX
XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX
XX WPI; 2000-533011/48.
XX
XX Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
XX Disclosure; Page 12; 121pp; English.
XX
XX The specification describes a beta-secretase enzyme. The enzyme cleaves

CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents an inhibitor of beta-secretase enzyme.
XX
SQ Sequence 8 AA;
Query Match 93.1%; Score 27; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VMXVAEF 7
DB 2 VMXVAEF 8
RESULT 3
AAY34479
ID AAY34479 standard; Protein; 821 AA.
XX AAY34479;
XX 25-AUG-1999 (first entry)
XX Porphorymonas gingivalis protein PG21.
XX
XX Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
KW vaccine; antigenic.
XX
XX Porphorymonas gingivalis.
XX
XX WO9929870-A1.
XX
XX 17-JUN-1999.
XX
XX 10-DEC-1998; 98WO-AU01023.
XX
XX 04-AUG-1998; 98AU-0005028.
XX
XX 10-DEC-1997; 97AU-0000839.
XX
XX 31-DEC-1997; 97AU-0001182.
XX
XX 30-JAN-1998; 98AU-0001546.
XX
XX 10-MAR-1998; 98AU-0002264.
XX
XX 09-APR-1998; 98AU-0002911.
XX
XX 23-APR-1998; 98AU-0003128.
XX
XX 05-MAY-1998; 98AU-0003338.
XX
XX 22-MAY-1998; 98AU-0003634.
XX
XX 29-JUL-1998; 98AU-0004917.
XX
XX (CSLC-) CSL LTD.
XX
XX Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
PI Ross BC, Rotnel LJ, Webb EA;
XX
XX WPI; 1999-385613/32.
XX
XX N-PSDB; AAX91697.
XX
XX Antigenic Porphorymonas gingivalis peptides for preventing
PT gingivitis
XX
XX Claim 1; Page 460-461; 588pp; English.
XX
XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphorymonas gingivalis. Probes can
CC be used to detect Porphorymonas gingivalis in standard hybridisation

CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis.

XX Sequence 821 AA;

Query Match 89.7%; Score 26; DB 20; Length 821.
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
| | | | |
Db 565 VMXVAEF 571

RESULT 4
AAV34354
ID AAV34354 standard; Protein; 869 AA.

XX AAV34354;

XX 25-AUG-1999 (first entry)

XX Porphyromonas gingivalis protein PG21.

XX Porphyromonas gingivalis; PG: periodontal disease; gingivitis;
KM vaccine; antigenic.

XX Porphyromonas gingivalis.

XX WO929870-A1.

XX 17-JUN-1999.

XX 10-DEC-1998; 98WO-AU01023.

XX 04-AUG-1998; 98AU-0005028.

XX 10-DEC-1997; 97AU-0000839.

XX 31-DEC-1997; 97AU-0001182.

XX 30-JAN-1998; 98AU-0001546.

XX 10-MAR-1998; 98AU-0002264.

XX 09-APR-1998; 98AU-0002911.

XX 23-APR-1998; 98AU-0003338.

XX 05-MAY-1998; 98AU-0003654.

XX 22-MAY-1998; 98AU-0004917.

XX (CSLC-) CSL LTD.

XX Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;

XX Ross BC, Rothel LJ, Webb EA;

XX MPI: 1999-385613/32.

XX N-PSDB; AAV31572.

XX Antigenic Porphyromonas gingivalis peptides for preventing

XX gingivitis

XX Claim 1: Page 316-318; 588pp; English.

XX AAV3156 to AAV31801 encode two hundred and sixty six antigenic

XX Porphyromonas gingivalis (PG) polypeptide sequences given in AAV34318 to

XX AAV4583. AAV31802 to AAV31989 represent PCR primers used in the

XX isolation of the PG polypeptides. The PG polypeptides have antibacterial

XX activity with a vaccine mechanism of action. The PG polypeptides can be

XX used as vaccines especially against Porphyromonas gingivalis. Probes can

XX be used to detect Porphyromonas gingivalis in standard hybridisation

XX assays. Porphyromonas gingivalis is involved in periodontal disease

XX especially gingivitis.

XX Sequence 869 AA;

XX Query Match 89.7%; Score 26; DB 20; Length 869;

XX Best Local Similarity 85.7%; Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
| | | | |
Db 613 VMXVAEF 619

RESULT 5
AAU56778
ID AAU56778 standard; Protein; 75 AA.

XX AAU56778;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #17674.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

XX dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX MPI: 2001-616774/71.

XX N-PSDB; AAS59578.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

XX vaccinating against and diagnosing infections, especially useful for

XX treating acne vulgaris -

XX Example 1; SEQ ID No 17973; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

XX polypeptides. The proteins and their associated DNA sequences are used in

XX the treatment, prevention and diagnosis of medical conditions caused by

XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

XX P. acnes is also involved in infections of bone, joints and the central

XX nervous system, however it is particularly involved in the inflammatory

XX lesions associated with acne vulgaris. A method for detecting the

XX presence or absence of P. acnes in a patient comprises contacting a

XX sample with a binding agent that binds to the proteins of the invention

XX and determining the amount of bound protein in the sample. The

XX polypeptides may be used as antigens in the production of antibodies

XX specific for P. acnes proteins. These antibodies can be used to

XX downregulate expression and activity of P. acnes polypeptides and

XX therefore treat P. acnes infections. The antibodies may also be used as

XX diagnostic agents for determining P. acnes presence, for example, by

XX enzyme linked immunosorbent assay (ELISA).

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 75 AA;

XX Query Match 82.8%; Score 24; DB 22; Length 75;

XX Best Local Similarity 71.4%; Pred. No. 56;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
11111
Db 1 VMACEF 7

RESULT 6

AAM41984
ID AAM41984 standard; Protein: 132 AA.

AC AAM41984;

DE 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6915.

XX Human; nootropic; immunosuppressant; cytosstatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokineic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB; AAI61140.

XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 6915; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytosstatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokineic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX Sequence 132 AA;
SQ

Query Match 82.8%; Score 24; DB 22; Length 132;

Best Local Similarity 71.4%; Pred. No. 1e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VMXVAEF 7
:11111
Db 64 LMVVAEF 70

RESULT 7

AAM41985
ID AAM41985 standard; Protein: 132 AA.

AC AAM41985;

DE 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6916.

XX Human; nootropic; immunosuppressant; cytosstatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokineic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB; AAI61141.

XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 6916; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytosstatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokineic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX Sequence 132 AA;
SQ

CC specification.
CC C.N.S disorders.
SQ Sequence 132 AA;
CC Note: The sequence data for this patent did not form part of the printed specification.

Query Match
Best Local Similarity 82.8%; Score 24; DB 22; Length 132;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
: | | | |
DB 64 LMVVAEF 70

RESULT 8
AAM41986
ID AAM41986 standard; Protein; 132 AA.
XX
AC AAM41986;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6917.
XX
KW Human; nocitropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
KW
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PE 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HSE-) HSEQ INC.
XX
PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX N-PSDB: AAI61142.
XX
DR Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6917; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nocitropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
SQ Sequence 132 AA;
CC Note: The sequence data for this patent did not form part of the printed specification.

Query Match
Best Local Similarity 82.8%; Score 24; DB 22; Length 132;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
: | | | |
DB 64 LMVVAEF 70

RESULT 9
AAV85820
ID AAV85820 standard; Protein; 225 AA.
XX
AC AAV85820;
XX
DT 10-APR-2000 (first entry)
XX
DE S. pneumoniae derived protein #29.
XX
KW Treatment; prevention; disease; diagnosis; gene therapy; screening; bacterial; antimicrobial; antibiotic; pathogenesis; infection.
KW
XX
OS Streptococcus pneumoniae.
XX
PN WO9806734-A1.
XX
PD 19-FEB-1998.
XX
PE 15-AUG-1997; 97WO-US14436.
XX
PR 16-AUG-1996; 96US-0024022.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
PI Stodola RK;
XX
XX WPI: 1998-159452/14.
XX N-PSDB: AAZ96199.
XX
DR Streptococcus pneumoniae proteins and related DNA - useful for
XX screening compounds for antibacterial activity
XX
XX Claim 5; Page 338-339; 640pp; English.
XX
XX This invention describes novel isolated Streptococcus pneumoniae
XX polynucleotides (see AAZ96173-296494) and their encoded proteins (see
XX AAZ85792-Y86182). The DNA, vectors and host cells described in the
XX method of the invention are useful for the recombinant expression of
XX polypeptides. The polypeptides are useful for treatment or prevention of
XX disease, or diagnosis of disease related to expression or activity of
XX such a polypeptide. They can also be used to screen for compounds which
XX interact with and inhibit or activate such a polypeptide. The
XX polypeptides (or DNA encoding them, via gene therapy) are also useful
XX for inducing an immunological response in a mammal. The antagonists are
XX useful to inhibit such bacterial polypeptides. The polypeptides are
XX particularly useful to identify antimicrobial compounds and antibiotics.
XX They are also useful to determine their role in pathogenesis of
XX infection, dysfunction and disease.
XX
XX Sequence 225 AA;
XX
XX Query Match
XX Best Local Similarity 82.8%; Score 24; DB 19; Length 225;
XX Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMAXAEF 7
11 :111
Db 115 VMSAEF 121

RESULT 10

AAU50802
ID AAU50802 standard; Protein: 236 AA.

AAU50802;

27-FEB-2002 (first entry)

Propionibacterium acnes immunogenic protein #11698.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN W0200181581-A2.

01-NOV-2001.

20-APR-2001; 2001WO-US12865.

21-APR-2000; 2000US-199047P.

02-JUN-2000; 2000US-208841P.

07-JUL-2000; 2000US-216747P.

(CORI-) CORIXA CORP.

Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

WPI: 2001-616774/71.

N-PSDB; AAS59549.

Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

Example 1; SEQ ID NO 11997; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence. For example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 236 AA;

Query Match 82.8%; Score 24; DB 22; Length 236;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMAXAEF 7
11 :111
Db 169 VMEIAEF 175

RESULT 11

AAV36930
ID AAV36930 standard; Protein: 254 AA.

AAV36930;

07-OCT-1999 (first entry)

Amino acid sequence of a Chlamydia trachomatis protein.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; peritrititis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.

OS Chlamydia trachomatis.

PN W09928475-A2.

10-JUN-1999.

27-NOV-1998; 98WO-IB01939.

04-NOV-1998; 98US-0107077.

28-NOV-1997; 97FR-0015041.

17-DEC-1997; 97FR-0016034.

(GEST) GENSET.

Griffals R;

WPI: 1999-371125/31.

Genome sequence of Chlamydia trachomatis

PT Disclosure; Page 786; 1755pp; English.

XX AAU36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, urethritis, epididymitis, genital diseases such as
CC nongonococcal urethritis, peritrititis, pneumopathy in breast feeding infants;
CC peritrititis, Bartholinitis; cervicitis, salpingitis,
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.

Sequence 254 AA;

Query Match 82.8%; Score 24; DB 20; Length 254;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMAXAEF 7
11 :111
Db 6 VMTVADF 12

RESULT 12

ABG18334
ID ABG18334 standard; Protein: 286 AA.

ABG18334;

18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #18325.
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB: AAS82521.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID NO 48693; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 286 AA:

Query Match 82.8%; Score 24; DB 22; Length 286;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
Db 23 VMQVAQF 29

RESULT 13
AAU37753
ID AAU37753 standard; Protein; 335 AA.
XX
AC AAU37753;
XX
DT 14-FEB-2002 (first entry)
XX
DE Streptococcus pneumoniae cellular proliferation protein #182.
XX

KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Streptococcus pneumoniae.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR N-PSDB: AAS55612.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID NO 13346; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 335 AA:

Query Match 82.8%; Score 24; DB 22; Length 335;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
|||
Db 224 VMSNAEF 230

RESULT 14
AAG20945
ID AAG20945 standard; Protein; 361 AA.
XX
AC AAG20945;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23321.
XX

KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
PN EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130892.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138440.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160816.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 82.8%; Score 24; DB 21; Length 361;
Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
11:111
Db 31 VMNMAEF 37

RESULT 15
AAG24458
ID AAG24458 standard; Protein; 361 AA.

AC AAG24458;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 28136.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hydrolisase assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.
PN

XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PF 05-MAR-1999; 99US-0123160.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134758.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139889.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0136624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145114.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 82.8%; Score 24; DB 21; Length 361;
Best Local Similarity 71.4%; Pred. No. 3,1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VMXVAEF 7
DB 31 VMNMAEF 37

Search completed: October 30, 2002, 12:27:07
Job time : 4.80344 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:23:30 ; Search time 1.34152 Seconds
(without alignments)
501.389 Million cell updates/sec

Title: US-09-724-571-78

Perfect score: 29

Sequence: 1 VMXVAEF 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :
1: PIR_71:*
2: PIR:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	96.6	2638	1 A42545	genome polyprotein
2	27	93.1	802	2 T05596	probable potassium
3	26	89.7	223	2 E64205	deoxyribose-phosph
4	26	89.7	224	2 S02216	deoxyribose-phosph
5	26	89.7	324	2 B84452	probable steroid s
6	25	86.2	201	2 F69988	hypothetical prote
7	25	86.2	632	2 T38126	probable electron
8	24	82.8	145	2 B41715	ribosomal protein
9	24	82.8	195	2 C97255	thymidine kinase (
10	24	82.8	236	2 B70728	hypothetical prote
11	24	82.8	248	2 T18315	hypothetical prote
12	24	82.8	249	2 G81693	serine/threonine p
13	24	82.8	335	2 F95103	6-phosphofructokin
14	24	82.8	335	2 D97971	6-phosphofructokin
15	24	82.8	367	2 F97010	fusion, chorismate
16	24	82.8	484	2 E86416	unknown protein, 3
17	24	82.8	565	2 B72660	probable type II D
18	24	82.8	582	2 E71052	hypothetical prote
19	24	82.8	584	2 F75090	archaeosine trna-r
20	24	82.8	754	2 T25551	hypothetical prote
21	24	82.8	919	2 T29581	hypothetical prote
22	24	82.8	921	2 H69643	isoleucine--trna 1
23	24	82.8	921	2 A83968	isoleucyl-trna syn
24	24	82.8	1122	2 T42400	Eph receptor tyros
25	23	79.3	89	2 T15018	hypothetical prote
26	23	79.3	96	2 AB3053	hypothetical prote
27	23	79.3	112	2 E90987	hypothetical prote
28	23	79.3	112	2 H85832	hypothetical prote
29	23	79.3	143	2 H69515	riboflavin synthas

30	23	79.3	157	2 G00016	FGF-receptor - com
31	23	79.3	162	2 G86842	shikimate kinase (
32	23	79.3	169	2 AB2739	acetyltransferase
33	23	79.3	172	2 S27019	fibroblast growth
34	23	79.3	172	2 S27020	fibroblast growth
35	23	79.3	172	2 S27021	fibroblast growth
36	23	79.3	172	2 S27022	fibroblast growth
37	23	79.3	187	2 S16314	photosynthetic rea
38	23	79.3	191	2 A82392	transcription regu
39	23	79.3	209	2 H97519	hypothetical prote
40	23	79.3	248	2 D81436	probable tomb tran
41	23	79.3	254	1 BVB953	phosphomannomutase
42	23	79.3	275	2 S28749	NADH dehydrogenase
43	23	79.3	291	2 G69479	methionyl aminopep
44	23	79.3	370	2 F36819	C14 protein - rabb
45	23	79.3	373	2 AC0253	ribonuclease III (

ALIGNMENTS

RESULT 1
A42545
genome polyprotein - Langat virus (strain TP21) (fragment)
N:Contains: nonstructural protein NS1; nonstructural protein NS2a; nonstructural prote
NS5
C:Species: Langat virus
C:Date: 30-Sep-1993 #sequence.revision 30-Sep-1993 #text.change 19-Jan-2001
C:Accession: A42545; A61409; G61409
R:Iacono-Connors, L.C.; Schmaljohn, C.S.
Virology 188, 875-880, 1992
A:Title: Cloning and sequence analysis of the genes encoding the nonstructural protei
A:Reference number: A42545; MUID:92263794
A:Accession: A42545
A:Molecule type: genomic RNA
A:Residues: 1-2638 <IAC>
A:Cross-references: GB:S35365; NID:g249315; PIDN:AAB22165.1; PID:g249316
R:Gutrahoo, F.; Helms, F.X.; Mandl, C.W.; Holzmann, H.; Kunz, C.; Gresikova, M.
J. Gen. Virol. 72, 333-338, 1991
A:Title: The relationship between the flaviviruses Skatlica and Langat as revealed by
A:Reference number: A61409; MUID:91132129
A:Accession: A61409
A>Status: not compared with conceptual translation
A:Molecule type: genomic RNA
A:Residues: 319-337 <GUI>
A:Accession: G61409
A>Status: not compared with conceptual translation
A:Molecule type: genomic RNA
A:Residues: 877-994 <GU2>
C:Superfamily: yellow fever virus genome polyprotein
C:Keywords: glycoprotein; nonstructural protein NS1 #status predicted <NS1>
F:353-562/Product: nonstructural protein NS1 #status predicted <NS1>
F:583-713/Product: nonstructural protein NS2a #status predicted <NS2a>
F:714-1334/Product: nonstructural protein NS2b #status predicted <NS2b>
F:912-919/Region: nucleotide-binding motif A (P-loop)
F:1335-1483/Product: nonstructural protein NS3 #status predicted <NS3>
F:1484-1735/Product: nonstructural protein NS4 #status predicted <NS4>
F:1736-2638/Product: nonstructural protein NS5 #status predicted <NS5>
F:85,207,223,873,1212,1671,1950/Binding site: carbohydrate (Asn) (covalent) #status p

Query Match 96.6%; Score 28; DB 1; Length 2638;
Best Local Similarity 85.7%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VMXVAEF 7
Db 153 VMTVAEF 159

RESULT 2
T05596
probable potassium transport protein F9D16.110 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 18-Aug-2000
 C:Accession: T05596
 R:Devan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hohsels, J.; Mewes, H.W.; Mayer, K.F.
 submitted to the Protein Sequence Database, February 1999
 A:Reference number: 215419
 A:Accession: T05596
 A:Molecule type: DNA
 A:Residues: 1-802 <BPV>
 A:Cross-references: EMBL:AL035394
 A:Experimental source: cultivar Columbia; BAC clone F9D16
 C:Genetics:
 A:Map position: 4
 A:Insertions: 1/3; 102/1; 185/1; 209/1; 296/1; 313/3; 352/1; 437/1
 A>Note: F9D16.110
 C:Superfamily: barley probable potassium transport protein HAK1
 C:Keywords: ion transport

Query Match 93.1%; Score 27; DB 2; Length 802;
 Best Local Similarity 71.4%; Pred. No. 43;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 ||:||||
 Db 629 VM5IAEF 635

RESULT 3

E64205
 deoxyribose-phosphate aldolase (EC 4.1.2.4) - Mycoplasma genitalium
 C:Species: Mycoplasma genitalium
 C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 21-Jul-2000
 C:Accession: E64205
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
 M.; Fuhmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
 C.A.; Venter, J.C.
 Science 270, 397-403, 1995
 A:Title: The minimal gene complement of Mycoplasma genitalium.
 A:Reference number: A64200; MUID:96026346
 A:Accession: E64205
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-223 <TIGR>
 A:Cross-references: GB:U93684; GB:L43967; NID:g3844650; PIDN:AAC71266.1; PID:g1045723;
 A:Experimental source: strain G-37
 C:Genetics:
 A:Genetic code: SGC3
 C:Superfamily: deoxyribose-phosphate aldolase
 C:Keywords: aldehyde-lyase; carbon-carbon lyase
 C:KeyWords: aldehyde-lyase; carbon-carbon lyase

Query Match 89.7%; Score 26; DB 2; Length 223;
 Best Local Similarity 71.4%; Pred. No. 21;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 ||:||||
 Db 93 VM5IAEF 99

RESULT 4

S02216
 deoxyribose-phosphate aldolase (EC 4.1.2.4) - Mycoplasma pneumoniae
 C:Species: Mycoplasma pneumoniae
 A:Varley, ATCC 29342
 C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 07-Dec-1999
 C:Accession: S02216; S73417
 R:Loechel, S.; Inamine, J.M.; Hu, P.C.
 Nucleic Acids Res. 17, 801, 1989
 A:Title: Nucleotide sequence of the deoc gene of Mycoplasma pneumoniae.
 A:Reference number: S02216; MUID:89128453
 A:Accession: S02216
 A:Molecule type: DNA
 A:Residues: 1-224 <LOE>

A:Cross-references: EMBL:X13544; NID:g44480; PIDN:CAA1897.1; PID:g44481
 R:Himmelfreid, R.; Hildert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumonia
 A:Reference number: S73327; MUID:97105885
 A:Accession: S73417
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-224 <HIM>
 A:Cross-references: EMBL:AE000011; GB:U00089; NID:g1673740; PIDN:AAB95739.1; PID:g167
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C:Genetics:
 A:Gene: deoc
 A:Genetic code: SGC3
 C:Superfamily: deoxyribose-phosphate aldolase
 C:Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 89.7%; Score 26; DB 2; Length 224;
 Best Local Similarity 71.4%; Pred. No. 22;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 ||:||||
 Db 93 VM5IAEF 99

RESULT 5

B84452
 probable steroid sulfotransferase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: B84452
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanNaken, S.E.; Umayam, L.; Tallon,
 eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: B84420; MUID:20083487
 A:Accession: B84452
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-324 <STO>
 A:Cross-references: GB:AE002093; NID:g4406768; PIDN:AAD20079.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g03770
 A:Map position: 2
 C:Superfamily: alcohol sulfotransferase

Query Match 89.7%; Score 26; DB 2; Length 324;
 Best Local Similarity 71.4%; Pred. No. 31;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 ||:||||
 Db 233 VM5IAEF 239

RESULT 6

F69988
 hypothetical protein ytbQ - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C:Accession: F69988
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.;
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Gal
 lecher, J.; Harwood, C.R.; Henaut, A.; Hildert, H.; Holstappel, S.; Hosono, S.; Hullo, M
 koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardino
 A.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scani

A:Authors: Schliech, S.; Schroeter, R.; Scoffone, F.; Sekiuchi, J.; Sekowska, A.; Serot
A:Author: M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033
A:Accession: F69988
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-201 <KUN>
A:Cross-references: GB:Z99119; GB:AL009126; NID:g2635411; PIDN:CAB14996.1; PID:e1185891;
A:Experimental source: strain 168
C:Genetics:
A:Gene: ytbQ

Query Match 86.2%; Score 25; DB 2; Length 201;
Best Local Similarity 57.1%; Pred. No. 36;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |:||||
Db 23 IMDIAEF 29

RESULT 7
T38126
Probable electron transfer flavoprotein precursor - fission yeast (*Schizosaccharomyces p*
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38126
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1997
A:Reference number: Z21772
A:Accession: T38126
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-632 <BAD>
A:Cross-references: EMBL:Z95334; PIDN:CAB08598.1; GSPDB:GN00066; SPDB:SPAC2068.04c
A:Experimental source: strain 972h-; cosmid c2068
C:Genetics:
A:Gene: SPDB:SPAC2068.04c
A:Map position: 1

Query Match 86.2%; Score 25; DB 2; Length 632;
Best Local Similarity 71.4%; Pred. No. 1,2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |:||||
Db 193 VMSIAEF 199

RESULT 8
B41715
ribosomal protein L13 [similarity] - *Halococcus marismortui*
C:Species: *Halococcus marismortui*
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: B41715
R:Kroemer, W.D.; Arndt, E.
J. Biol. Chem. 266, 24573-24579, 1991
A:Title: Halobacterial S9 operon. Three ribosomal protein genes are cotranscribed with g
bacterial) marismortui.
A:Reference number: A41715; MUID:92105119
A:Accession: B41715
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-145 <KRO>
A:Cross-references: GB:M76567; NID:g148775; PIDN:AAA73097.1; PID:g148777
C:Superfamily: Escherichia coli ribosomal protein L13

Query Match 82.8%; Score 24; DB 1; Length 145;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
 |:||||
Db 1 MCVAEF 6

RESULT 9
C97255
thymidine kinase (EC 2.7.1.21) [similarity] - *Clostridium acetobutylicum*
C:Species: *Clostridium acetobutylicum*
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 31-Dec-2001
C:Accession: C97255
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97255
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-195 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK0830.1; PID:g15025935; GSPDB:GN00168
A:Experimental source: *Clostridium acetobutylicum* ATCC824
C:Genetics:
A:Gene: CAC287
C:Superfamily: thymidine kinase
C:Keywords: phosphotransferase

Query Match 82.8%; Score 24; DB 2; Length 195;
Best Local Similarity 57.1%; Pred. No. 65;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |:||||
Db 131 LMAIAEF 137

RESULT 10
B70728
hypothetical protein RV2558 - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70728
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: B70728
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <COL>
A:Cross-references: GB:Z77250; GB:AL123456; NID:g3261617; PIDN:CAB01046.1; PID:e25533
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2558

Query Match 82.8%; Score 24; DB 2; Length 236;
Best Local Similarity 71.4%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
 |:||||
Db 217 VLDVAEF 223

RESULT 11
T18315
hypothetical protein L7610.5 - *Leishmania major*
C:Species: *Leishmania major*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T18315
 R:Oliver, K.; Murphy, L.; Quail, M.; Lawson, D.; Harris, D.; Rajandream, M.; Ivens, A.;
 submitted to the EMBL Data Library, May 1999
 A:Reference number: 218876
 A:Accession: T18315
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-248 <OLJ>
 A:Cross-references: EMBL:AL034356; NID:e1371878; PID:e1371560; PIDN:CAA2241.1
 C:Genetics:
 A:Note: U7610.5

Query Match 82.8%; Score 24; DB 2; Length 248;
 Best Local Similarity 71.4%; Pred. No. 83;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 || ||||
 Db 196 LMRVAEF 202

RESULT 12

G81693
 serine/threonine protein phosphatase, probable TC0530 [imported] - Chlamydia muridarum
 C:Species: Chlamydia muridarum, Chlamydia trachomatis Mopn
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
 C:Accession: G81693
 R:Read, T.D.; Brunham, R.C.; Shen, C.; GILL, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
 A:Reference number: A81500; MUID:20150255
 A:Accession: G81693
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-249 <TEU>
 A:Cross-references: GB:AE002321; GB:AE002160; NID:g7190560; PIDN:AAE9372.1; PID:g719057
 A:Experimental source: strain Nigg (Mopn)
 C:Genetics:
 A:Gene: TC0530
 C:Superfamily: conserved hypothetical protein y100; conserved hypothetical protein y100

Query Match 82.8%; Score 24; DB 2; Length 249;
 Best Local Similarity 71.4%; Pred. No. 84;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 || ||||
 Db 1 MMIVAEF 7

RESULT 13

F95103
 6-phosphofructokinase [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: F95103
 R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held,
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: F95103
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-335 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75023.1; PID:g1497371; GSPDB:GN00164; TIGR:SP4
 C:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SFO896
 C:Superfamily: 6-phosphofructokinase; 6-phosphofructokinase 1 homology

Query Match 82.8%; Score 24; DB 2; Length 335;
 Best Local Similarity 71.4%; Pred. No. 11e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 || ||||
 Db 224 VMSAAEF 230

RESULT 14

D97971
 6-phosphofructokinase (EC 2.7.1.11) [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C:Accession: D97971
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: D97971
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-335 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAK9600.1; PID:g15458395; GSPDB:GN00174
 C:Genetics:
 A:Gene: pfkA
 C:Superfamily: 6-phosphofructokinase; 6-phosphofructokinase 1 homology
 C:Keywords: phosphotransferase

Query Match 82.8%; Score 24; DB 2; Length 335;
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 || ||||
 Db 224 VMSAAEF 230

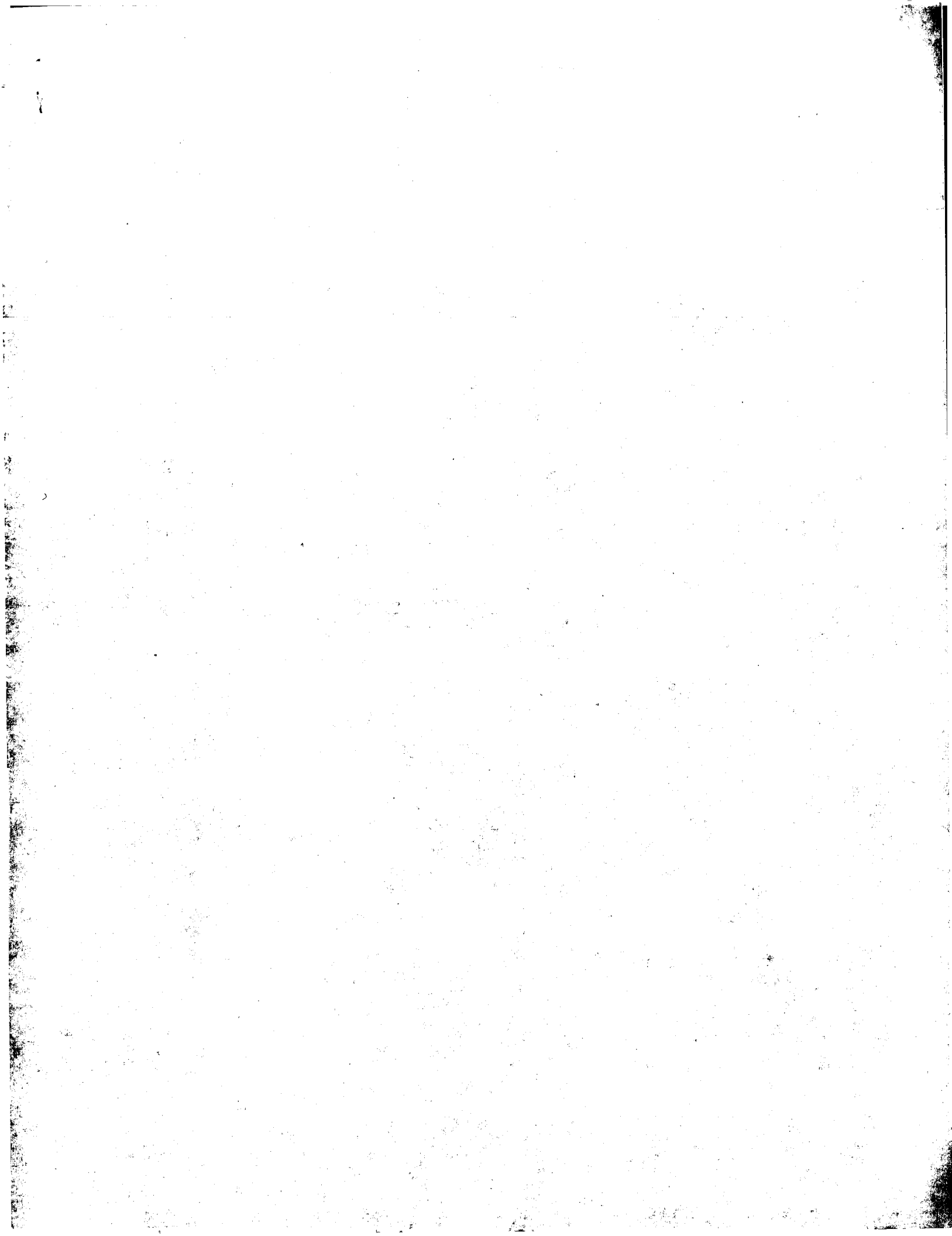
RESULT 15

F97010
 fusion, chorismate mutase and splikimate 5-dehydrogenase [imported] - Clostridium acet
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C:Accession: F97010
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: F97010
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-367 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK78873.1; PID:g15023795; GSPDB:GN00168
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC0897

Query Match 82.8%; Score 24; DB 2; Length 367;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 || ||||
 Db 29 VMEVAEF 35

Search completed: October 30, 2002, 12:31:31
 Job time : 4.34152 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:10 ; Search time 0.601966 Seconds
(without alignments)
450.253 Million cell updates/sec

Title: US-09-724-571-78
Perfect score: 29
Sequence: 1 VMXVAEF 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	96.6	3414	1	P29837 1 genome po
2	26	89.7	223	1	P47296 mycoplasma
3	26	89.7	224	1	P09924 mycoplasma
4	25	86.2	253	1	P53560 bacillus su
5	25	86.2	632	1	P87111 s probable
6	24	82.8	145	1	P29198 halocarcula
7	24	82.8	157	1	P29198 halocarcula
8	24	82.8	236	1	P29198 halocarcula
9	24	82.8	565	1	P29198 halocarcula
10	24	82.8	921	1	P29198 halocarcula
11	24	82.8	971	1	P29198 halocarcula
12	24	82.8	1095	1	P29198 halocarcula
13	23	79.3	143	1	P29198 halocarcula
14	23	79.3	162	1	P29198 halocarcula
15	23	79.3	187	1	P29198 halocarcula
16	23	79.3	254	1	P29198 halocarcula
17	23	79.3	275	1	P29198 halocarcula
18	23	79.3	291	1	P29198 halocarcula
19	23	79.3	299	1	P29198 halocarcula
20	23	79.3	332	1	P29198 halocarcula
21	23	79.3	632	1	P29198 halocarcula
22	23	79.3	654	1	P29198 halocarcula
23	23	79.3	773	1	P29198 halocarcula
24	23	79.3	773	1	P29198 halocarcula
25	23	79.3	801	1	P29198 halocarcula
26	23	79.3	802	1	P29198 halocarcula
27	23	79.3	806	1	P29198 halocarcula
28	23	79.3	806	1	P29198 halocarcula
29	23	79.3	808	1	P29198 halocarcula
30	23	79.3	813	1	P29198 halocarcula
31	23	79.3	819	1	P29198 halocarcula
32	23	79.3	821	1	P29198 halocarcula
33	23	79.3	821	1	P29198 halocarcula

34	23	79.3	822	1	FGRL_HUMAN	P11362 homo sapien
35	23	79.3	822	1	FGRL_MOUSE	P16092 mus musculu
36	23	79.3	822	1	FGRL_RAT	P04589 ratu
37	23	79.3	823	1	CEK3_CHICK	P18461 ictaluri
38	23	79.3	852	1	VTER_HSV1	P00140 ictaluri
39	23	79.3	980	1	POLG_LIV	P22338 loup
40	23	79.3	1002	1	EPB5_CHICK	P07497 gallu
41	23	79.3	1052	1	FGRL_DROME	P09147 drosophila
42	23	79.3	1914	1	STCK_EMENT	P00706 emeritella
43	23	79.3	3412	1	POLG_TBREV	P07720 t genome po
44	23	79.3	3414	1	POLG_TBREV	P01299 t genome po
45	23	79.3	3414	1	POLG_TBREV	P14336 t genome po

ALIGNMENTS

RESULT 1
ID POLG_LANVT STANDARD: PRT: 3414 AA.
AC P29637.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide (Contains: Capsid protein C (core protein);
DE Envelope protein PMV; Matrix protein (Envelope protein M); Major
DE envelope protein E; Nonstructural protein NS1; Nonstructural protein
DE NS2A; Nonstructural protein NS2B; Helicase/protease (EC 3.4.21.98)
DE (NS3); Nonstructural protein NS4A; Nonstructural protein NS4B; RNA-
DE directed RNA polymerase (EC 2.7.7.48) (NS5)).
DE Langat virus (strain TP21).
OS Viruses: ssRNA positive-strand viruses, no DNA stage: Flaviviridae;
OC
NCBI_TaxID=31638;
RN
RN
RP SEQUENCE OF 1-776 FROM N.A.
RX MEDLINE=92074260; PubMed=1720591;
RA Mandl C.W., Iacono-Connors L., Wallner G., Holzmann H., Kunz C.,
RA Heinz F.X.;
RT "Sequence of the genes encoding the structural proteins of the low-
RT virulence tick-borne flaviviruses Langat TP21 and Yelantsev.";
RL Virology 185:891-895(1991).
[2]
SEQUENCE OF 777-3414 FROM N.A.
RX MEDLINE=92263794; PubMed=1316684;
RA Iacono-Connors L.C., Schmaljohn C.S.;
RT "Cloning and sequence analysis of the genes encoding the
RT nonstructural proteins of Langat virus and comparative analysis with
RT other flaviviruses.";
RL Virology 188:875-880(1992).
CC
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF FOUR PEPTIDE BONDS IN THE VIRAL
CC PRECURSOR POLYPEPTIDE, COMMONLY WITH ASP OR GLU IN THE P6
CC POSITION, CYS OR THR IN P1 AND SER OR ALA IN P1.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb.ch).
CC
CC EMBL: M73835; AAA02740.1; ALT_TERM.
CC EMBL: S35365; AAB22165.1; -.
CC PIR: A41704; A41704.
CC PIR: A42545; A42545.
CC HSSP: P14336; 1SVB.
CC MEROPS: S07.001; -.
CC InterPro: IPR001410; DEAD.

```

DR InterPro: IPR000069; Flavl_M.
DR InterPro: IPR001157; Flavl_NSI.
DR InterPro: IPR000752; Flavl_NS2A.
DR InterPro: IPR000487; Flavl_NS2B.
DR InterPro: IPR000404; Flavl_NS4A.
DR InterPro: IPR001528; Flavl_NS4B.
DR InterPro: IPR000208; Flavl_NS5.
DR InterPro: IPR001122; Flavl_capsid.
DR InterPro: IPR000336; Flavl_glycoprote.
DR InterPro: IPR001850; Flavl_helicase.
DR InterPro: IPR002535; Flavl_propep.
DR InterPro: IPR002877; FtsJ.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF01003; Flavl_capsid; 1.
DR Pfam: PF00869; Flavl_glycoprot; 1.
DR Pfam: PF02832; Flavl_glycop_C; 1.
DR Pfam: PF00949; Flavl_helicase; 1.
DR Pfam: PF01004; Flavl_M; 1.
DR Pfam: PF00948; Flavl_NSI; 1.
DR Pfam: PF01005; Flavl_NS2A; 1.
DR Pfam: PF01002; Flavl_NS2B; 1.
DR Pfam: PF01350; Flavl_NS4A; 1.
DR Pfam: PF01349; Flavl_NS4B; 1.
DR Pfam: PF00972; Flavl_NS5; 1.
DR Pfam: PF01570; Flavl_propep; 1.
DR Pfam: PF01728; FtsJ; 1.
DR Pfam: PF00271; helicase_C; 1.
DR Pfam: PF001496; Flavl_NSI; 1.
DR Pfam: PF001556; Flavl_glycoprote; 1.
DR SMART; SMO0490; HELIC; 1.
DR Polyprotein; Glycoprotein; Transferrase; RNA-directed RNA polymerase;
KM Core protein; Envelope protein; Hydrolase; Helicase;
KW ATP-binding; Transmembrane; Nonstructural protein
FT INIT_MBT 1
FT CHAIN 1 112
FT CHAIN 113 205
FT CHAIN 206 280
FT CHAIN 281 776
FT CHAIN 777 1128
FT CHAIN 1129 1358
FT CHAIN 1359 1489
FT CHAIN 1490 2110
FT CHAIN 2111 2259
FT CHAIN 2260 2511
FT CHAIN 2512 3414
FT NP_BIND 1688 1695
FT SITE 1779 1782
FT TRANSMEM 103 119
FT TRANSMEM 262 278
FT TRANSMEM 728 744
FT TRANSMEM 758 774
FT TRANSMEM 283 310
FT DISULFID 340 396
FT DISULFID 354 385
FT DISULFID 372 401
FT DISULFID 466 570
FT DISULFID 587 618
FT CARBOHYD 144 144
FT CARBOHYD 434 434
SQ SEQUENCE 3414 AA; 378017 MW; 59CBFE95DD70D82E CRC64;

```

```

Query Match 96.6%; Score 28; DB 1; Length 3414;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 VMXVAEF 7
| | | | |
DB 929 VMTVAEF 935

```

RESULT 2

```

DEOC_MYCGE
ID DEOC_MYCGE STANDARD: PRT; 223 AA.
AC P47296;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Deoxyribose-phosphate aldolase (EC 4.1.1.2.4) (Phosphodeoxyribosaldolase)
DE (Deoxyribosaldolase).
GN DEOC OR MG050.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -1- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 5-phosphate = D-
CC glyceraldehyde 3-phosphate + acetaldehyde.
CC -1- PATHWAY: NUCLEOTIDE AND DEOXYRIBONUCLEOTIDE CATABOLISM.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES.
CC DEOC SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U39684; AAC71266.1; -.
CC TIGR: MG050; -.
DR InterPro: IPR002915; DEOC.
DR InterPro: IPR003009; FNN_enzyme.
DR Pfam: PF01791; DEOC; 1.
KM Lyase; Schiff base; Complete proteome.
FT BINDING 152 152
FT BINDING 152 152
SQ SEQUENCE 223 AA; 24675 MW; 332430231CE99DB0 CRC64;

```

```

Query Match 89.7%; Score 26; DB 1; Length 223;
Best Local Similarity 71.4%; Pred. No. 8;7;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 VMXVAEF 7
| | | | |
DB 93 VMTVAEF 99

```

```

RESULT 3
DEOC_MYCPN
ID DEOC_MYCPN STANDARD: PRT; 224 AA.
AC P09924;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Deoxyribose-phosphate aldolase (EC 4.1.1.2.4) (Phosphodeoxyribosaldolase)
DE (Deoxyribosaldolase).
GN DEOC OR MPN063 OR MP091.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]

```

RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 29342 / M129;
 RX MEDLINE=89128453; PubMed=2492658;
 RA Loechel S., Inamine J.M., Hu P.-C.;
 RT "Nucleotide sequence of the deoc gene of Mycoplasma pneumoniae";
 RL Nucleic Acids Res. 17:801-801(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelfreuch R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
 RT Herrmann R.;
 RL "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae";
 CC Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 5-phosphate - D-
 CC -1- glyceraldehyde 3-phosphate + acetaldehyde.
 CC -1- PATHWAY: NUCLEOTIDE AND DEOXYRIBONUCLEOTIDE CATABOLISM.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES.
 CC DEOC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X13544; CAA31897.1; -
 CC DR EMBL: AE000011; AAB95739.1; -
 CC DR PIR: S02216; S02216.
 CC DR InterPro: IPR002915; Deco.
 CC DR Pfam: PF01791; Deco; 1.
 CC KW Lyase; Schiff base; Complete proteome.
 CC FT BINDING 152 152 SCHIFF-BASE (BY SIMILARITY).
 CC SEQUENCE 224 AA; 24878 MW; 73CCE4932E7881F7 CRC64;
 SO
 Query Match 89.7%; Score 26; DB 1; Length 224;
 Best Local Similarity 71.4%; Pred. No. 8; Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 VNMVAEF 7
 Db 93 VNMIAEF 99
 RESULT 4
 YTBQ_BACSU
 ID YTBQ_BACSU STANDARD: PRT: 253 AA.
 AC P53560;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ytbQ.
 GN YTBQ.
 OS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 CC NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96312354; PubMed=8763940;
 RA Bower S., Perkins J.B., Yocum R.R., Howitt C.L., Rahaim P.,
 RT Peto J.;
 RL "Cloning, sequencing, and characterization of the Bacillus subtilis
 RL biotin biosynthetic operon";
 CC J. Bacteriol. 178:4122-4130(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U51868; AAB17463.1; -
 CC DR Subtilist; BG11787; ytbQ.
 CC KW Hypothetical protein; Complete proteome.
 CC SEQUENCE 253 AA; 28233 MW; 572CC4681F04C650 CRC64;
 SO
 Query Match 86.2%; Score 25; DB 1; Length 253;
 Best Local Similarity 57.1%; Pred. No. 18; Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 VNMVAEF 7
 Db 75 IMDIAEF 81
 RESULT 5
 ETFD_SCHPO
 ID ETFD_SCHPO STANDARD: PRT: 632 AA.
 AC P87111;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Probable electron transfer flavoprotein-ubiquinone oxidoreductase,
 DE mitochondrial precursor (EC 1.5.5.1) (ETF-00) (ETF-ubiquinone
 DE oxidoreductase) (ETF dehydrogenase) (Electron-transferring-
 DE flavoprotein dehydrogenase).
 GN SPAC2068.04C.
 OS Schizosaccharomyces pombe (fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RL Badcock K., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
 RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ACCEPTS ELECTRONS FROM ETF AND REDUCES UBIQUINONE (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Reduced ETF + ubiquinone -> ETF + ubiquinol.
 CC -1- COFACTOR: FAD AND A 4FE-4S CLUSTER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE ETF-00 / FIXC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z95334; CAB08598.1; -
 CC DR Oxidoreductase; Electron transport; Flavoprotein; FAD; Iron-sulfur;
 CC KW 4Fe-4S; Mitochondrion; Transist peptide; Ubiquinone.
 CC FT TRANSIT 1 ?
 CC FT CHAIN ? 632
 CC FT NE_BIND 93 107
 CC FT METAL 575 575
 CC FT METAL 601 601
 CC FT METAL 604 604
 CC FT METAL 607 607
 CC SEQUENCE 632 AA; 69472 MW; 1B0F22374E33771B CRC64;
 SO
 Query Match 86.2%; Score 25; DB 1; Length 632;
 Best Local Similarity 71.4%; Pred. No. 45; Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXAEF 7
 11111
 Db 193 VMSIAEF 199

RESULT 6
 RL13_HALMA STANDARD; PRT: 145 AA.
 AC P29198;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 50S ribosomal protein L13p (Hmal13).
 GN RPL13p.
 OS Halorcula marismortui (Halobacterium marismortui).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloarcula.
 CC NCBI_TaxID=2238;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92105119; PubMed=1840597;
 RA Kroemer, W.J., Arndt E.;
 RT "Halobacterial 5S operon. Three ribosomal protein genes are
 RT cotranscribed with genes encoding a tRNA(Leu), the enolase, and a
 RT putative membrane protein in the archaeobacterium Haloarcula
 RT (Halobacterium) marismortui.";
 RL J. Biol. Chem. 266:24573-24579(1991).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RC STRAIN=ATCC 43049; PubMed=10937989;
 RX MEDLINE=20396344; PubMed=10937989;
 RA Ban N., Nissen P., Hansen J., Moore P.B., Steitz T.A.;
 RT "The complete atomic structure of the large ribosomal subunit at 2.4
 RT A resolution.";
 RL Science 289:905-920(2000).
 CC -1- SIMILARITY: BELONGS TO THE L13P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M76567; AAA73097.1; -
 DR PIR: B41715; B41715.
 DR PDB: 1FFK; 14-AUG-00.
 DR InterPro: IPR001074; Ribosomal_L13.
 DR Pfam: PF00572; Ribosomal_L13.
 DR ProDom: PD001791; Ribosomal_L13; 1.
 DR ProSite: PS00783; RIBOSOMAL_L13; 1.
 KW Ribosomal protein; 3D-structure.
 SO SEQUENCE 145 AA; 16228 MW; 069CE66662AE3BC CRC64;

Query Match 82.8%; Score 24; DB 1; Length 145;
 Best Local Similarity 83.3%; Pred. No. 20;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 MXVAEF 7
 11111
 Db 1 MSVAEF 6

RESULT 7
 RAP_TAROF STANDARD; PRT: 157 AA.
 ID RAP_TAROF
 AC 049065;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Root allergen protein (RAP).
 OS Taraxacum officinale (Common dandelion).

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asterales; euasterids II; Asterales; Asteraceae; Lactuceae;
 CC Taraxacum.
 CC NCBI_TaxID=50225;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Root;
 RA Xu X.-Y., Bewley J.D., Greenwood J.S.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE BETV1 FAMILY OF PATHOGENESIS-RELATED
 CC PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF036931; AAB92255.1; -
 DR HSSP: P15494; 1BVL.
 DR Mendel: 28037; Tarof.1173;28037.
 DR InterPro: IPR000916; Bet_v_1.
 DR Pfam: PF00407; Bet_v_1; 1.
 DR PRINTS: PR00634; BETALLERGEN.
 DR ProDom: PD000531; Bet_v_1; 1.
 DR ProSite: PS00451; PATHOGENESIS_BETV1; 1.
 KW Allergen.
 SO SEQUENCE 157 AA; 17040 MW; 5892AB8593A8A7E0 CRC64;

Query Match 82.8%; Score 24; DB 1; Length 157;
 Best Local Similarity 83.3%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 MXVAEF 7
 11111
 Db 1 MAVAEF 6

RESULT 8
 YP58_MYCTU STANDARD; PRT: 236 AA.
 ID YP58_MYCTU
 AC 050740;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 25.7 kDa protein RV2558.
 GN RV2558 OR MT2635 OR MTCY9C4.10C.
 OS Mycobacterium tuberculosis.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Whitehead B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RA complete genome sequence.";
 RT Nature 393:537-544(1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV2557.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: Z77250; CAB01046.1; -
DR EMBL: AE007098; AKK46947.1; -
DR TIGR: MT635; -
DR TubercuList: RV2558; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 236 AA; 25718 MW; 13E3B049DBF79C6B CRC64;

Query Match 82.8%; Score 24; DB 1; Length 236;
Best Local Similarity 71.4%; Pred. No. 32;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAF 7
I: ||||
DB 217 VLDVAF 223

RESULT 9
TP6B_AERPE
ID TP6B_AERPE STANDARD: PRT; 565 AA.
AC O9Y64;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Type II DNA topoisomerase VI subunit B (EC 5.99.1.3).
GN TOPEB OR APE0706.
OS Aeropyrum pernix.
OC Archaea; Ctenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jinno K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1";
RL DNA Res. 6:83-101(1999).
CC -1- FUNCTION: RELAXES BOTH POSITIVE AND NEGATIVE SUPERURNS AND
EXHIBITS A STRONG DECATENASE ACTIVITY. THE B SUBUNIT BINDS ATP (BY
SIMILARITY)
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
of double-stranded DNA.
CC -1- SUBUNIT: HETEROTETRAMER OF TWO SUBUNITS A AND TWO SUBUNITS B (BY
SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000060; BAA/9682.1; -
DR InterPro: IPR003594; HATPase_C.
DR InterPro: IPR004359; HIS_KIN_sig.
DR Pfam: PF02518; HATPase_C; 1.
DR SMART: SM00387; HATPase_C; 1.
KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Complete proteome.
SQ SEQUENCE 565 AA; 64096 MW; E4A3DFC5E53D508 CRC64;

Query Match 82.8%; Score 24; DB 1; Length 565;
Best Local Similarity 83.3%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 MXYAER 7
I: ||||
DB 28 MSVAF 33

RESULT 10
SYL_BACSU
ID SYL_BACSU STANDARD: PRT; 921 AA.
AC Q45477; P71022; O31730;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Isoleucyl-tRNA synthetase (EC 6.1.1.5) (Isoleucine--tRNA ligase)
DE (IleRS).
GN ILES.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RP [2]
RC SEQUENCE OF 1-121 FROM N.A.
RA Stewart G.C., Cha J.H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RP [3]
RP SEQUENCE OF 733-921 FROM N.A.
RC STRAIN-168;
RA Pragai Z., Tjalsma H., Bolhuis A., van Dijk J.M., Venema G.,
RA Bron S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-Isoleucine + tRNA(Ile) = AMP +
diphosphate + L-Isoleucyl-tRNA(Ile).
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: Z69112; CAB1417.1; -
DR EMBL: U60901; AAB49280.1; -
DR EMBL: U48870; AAB57764.1; -
DR HSSP: P41972; 1003.
DR Subtilist: BG1792; ILES.
DR InterPro: IPR002300; tRNA-synt_1a.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002301; tRNA-synt_1le.
DR Pfam: PF00133; tRNA-synt_1; 1.

```

DR PRINTS: PR00984; TRNASTYHILE.
DR PROSITE; PS00178; AA-RNA-LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Zinc; Complete proteome.
FT SITE 57 67 "HIGH" REGION.
FT BINDING 594 598 ATP (BY SIMILARITY).
FT CONFLICT 744 744 S -> F (IN REF. 3).
SQ SEQUENCE 921 AA; 104784 MW; 011F31E5F7460D43 CRC64;

Query Match 82.8%; Score 24; DB 1; Length 921;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 MXVAEF 7
DB 116 MSVAEF 121

RESULT 11
AMPN_HAECC STANDARD; PRT; 971 AA.
ID AMPN_HAECC
AC 010737;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aminoacyl-tRNA synthetase (Membrane
glycoprotein H11).
OS Haemonchus contortus (Barber pole worm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
OX NCBI_Taxid=6289;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-927373974; PubMed-9128148;
RA Smith T.S., Graham M., Munn E.A., Newton S.E., Knox D.P.,
RA Coatswell W.J., McMichael-Phillips D., Smith H., Smith W.D.,
RA Oliver J.J.;
RT "Cloning and characterization of a microsomal aminoacyl-tRNA synthetase from the
RT intestine of the nematode Haemonchus contortus."
RL Biochim. Biophys. Acta 1338:295-306(1997).
CC -1- CATALYTIC ACTIVITY: Release of an N-terminal amino acid,
CC preferably a neutral or hydrophobic one, from a polypeptide.
CC Aminoacyl-arylamides are poor substrates.
CC -1- COFACTOR: BINDS 1 ZINC ION.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X94187; CAA63897.1; -
DR MEROPS; M01.001; -
DR InterPro; IPR001930; Aladiptase.
DR InterPro; IPR00130; Zn_MTPeptidase.
DR Pfam; PF01433; Peptidase_M1; 1.
DR PRINTS; PR00756; ALADIPTASE.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolyase; Metalloprotease; Aminoacyl-tRNA synthetase; Zinc; Glycoprotein;
KW Transmembrane; Signal-anchor.
FT INIT_MET 0 0 BY SIMILARITY.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 38 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT TRANSMEM 17 38 (POTENTIAL).
FT DOMAIN 39 971 EXTRACELLULAR (POTENTIAL).
FT METAL 378 378 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 379 379 BY SIMILARITY.
FT METAL 382 382 ZINC (CATALYTIC) (BY SIMILARITY).

```

```

FT METAL 401 401 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 465 465 PROTON DONOR (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 548 548 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 971 AA; 110542 MW; 95C6A92B5CCA227C CRC64;

Query Match 82.8%; Score 24; DB 1; Length 971;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
DB 541 VISVAEF 547

RESULT 12
AT9B_MOUSE STANDARD; PRT; 1095 AA.
ID AT9B_MOUSE
AC P98195;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Potential phospholipid-transporting ATPase IIb (EC 3.6.3.1).
GN AT9B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE-20473714; PubMed-11015572;
RA Halleck M.S., Lawler J.F. Jr., Blackshaw S., Gao L., Nagarajan P.,
RA Hacker C., Pyle S., Newman J.T., Nakenshi Y., Ando H., Weinstock D.,
RA Williamson P., Schlegel R.A.;
RT "Differential expression of putative transmembrane amphipath
RT transporters."
RL Physiol. Genomics 1:139-150(1999).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: FOUND IN MOST TISSUES EXCEPT SPLEEN AND
CC MUSCLE. MOST ABUNDANT IN TESTIS. ALSO DETECTED IN FETAL TISSUES.
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (EI-E2
CC ATPASES). SUBFAMILY IV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF155913; AAF08476.1; -
DR MGD; MGI:1354757; At9B.
DR InterPro; IPR001757; EI-E2-ATPase.
DR InterPro; IPR001454; Hydrolyase.
DR Pfam; PF00122; EI-E2-ATPase; 1.
DR Pfam; PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CATALPASE.
DR PROSITE; PS00154; ATPASE_EI_E2; 1.
KW Hydrolyase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
KW Multigene family.
FT DOMAIN 1 94 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 95 116 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 117 121 POTENTIAL.
FT TRANSMEM 122 144 POTENTIAL.
FT TRANSMEM 145 328 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 329 349 POTENTIAL.
FT DOMAIN 350 357 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 358 379 POTENTIAL.

```

```

FT DOMAIN 380 878 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 879 899 POTENTIAL.
FT DOMAIN 900 911 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 912 930 POTENTIAL.
FT DOMAIN 931 960 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 961 979 POTENTIAL.
FT DOMAIN 980 986 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 987 1009 POTENTIAL.
FT DOMAIN 1010 1015 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 1016 1036 POTENTIAL.
FT DOMAIN 1037 1053 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 1054 1078 POTENTIAL.
FT DOMAIN 1079 1095 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 416 416 PHOSPHORYLATION (BY SIMILARITY).
FT METAL 822 822 MAGNESIUM (BY SIMILARITY).
FT METAL 826 826 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 1095 AA; 123506 MW; 76D37684241D176E CRC64;

Query Match
Best Local Similarity 82.8%; Score 24; DB 1; Length 1095;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVXAEF 7
DB 1016 LMVAAEF 1022

RESULT 13
RIBB_ARCFU STANDARD: PRT: 143 AA.
AC 028152;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 41, Last annotation update)
DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)
GN RIBH OR RIBE OR AF2128.
OS Archaeoglobus fulgidus.
OC Archaea: Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Richardson D.L., Gwinn M., Hickey E.K., Peterson J.D.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sulten G.G., Gill S.,
RA Kiknes E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Goeysse J.D., Weidman J.F., McDonald L., Uitterback T.,
RA Cotton M.D., Spriggs T., Artlach P., Kalne B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- FUNCTION: Riboflavin synthase is a bifunctional enzyme complex
catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
ribityl-amino-2,4(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-
butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit
catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-
2,4(1H,3H)-pyrimidinone with L-3,4-dihydroxy-2-butanone-4-
phosphate yielding 6,7-dimethyl-8-lumazine (By similarity).
CC -1- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =
riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
CC -1- PATHWAY: Riboflavin synthesis; last step.
CC -1- SIMILARITY: BELONGS TO THE DMRL SYNTHASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000957; AAB89124.1; -.
DR HSSP: P11998; IRVV.
DR TIGR: AF2128; -.
DR InterPro: IPR002180; DMRL_synthase.
DR Pfam: PF00885; DMRL_synthase.1.
KW Riboflavin biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 143 AA; 15645 MW; BBA3141206905410 CRC64;

Query Match
Best Local Similarity 79.3%; Score 23; DB 1; Length 143;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MVXAEF 7
DB 8 MVXAEF 13

RESULT 14
AROK_LACLA STANDARD: PRT: 162 AA.
ID AROK_LACLA
AC 09CEU1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Shikimate kinase (EC 2.7.1.71) (SK).
GN AROK OR IL1743.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Boletín A., Wincker P., Manger S., Jallón O., Malarre K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -1- CATALYTIC ACTIVITY: ATP + shikimate = ADP + shikimate 3-phosphate.
CC -1- PATHWAY: FIFTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE SHIKIMATE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB006404; AAK05841.1; -.
DR InterPro: IPR000623; Shik_kinase.
DR Pfam: PF01202; SKI.1.
DR PRINTS: PR01100; SHIKIMKINASE.
DR PROSITE: PS01128; SHIKIMATE_KINASE.1.
KW Aromatic amino acid biosynthesis; Transferase; Kinase; ATP-binding;
KW Complete proteome.
FT NP_BIND 7 14 ATP (POTENTIAL).
SQ SEQUENCE 162 AA; 18383 MW; 1F7FDEAF42AE9039 CRC64;

Query Match
Best Local Similarity 79.3%; Score 23; DB 1; Length 162;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 2 MXVAEF 7
1:111
Db 37 MTIAEF 42

RESULT 15

CYCR_ERVSP STANDARD; PRT; 187 AA.

AC P26278; 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Photosynthetic reaction center cytochrome C subunit (Fragment).

CYTC Erythrobacter sp. (strain OCH 114).

OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;

OC Roseobacter.

OX NCBI_TaxID=2434;

RN 11

RP SEQUENCE FROM N.A. MEDLINE=92157872; PubMed=1787796;

RA Liebetanz R., Hornberger U., Drews G.;

RT "Organization of the genes coding for the reaction-centre L and M

subunits and B870 antenna polypeptides alpha and beta from the

aerobic photosynthetic bacterium Erythrobacter species OCH114.";

RL Mol. Microbiol. 5:1459-1468(1991).

-1- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A

TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO

OXIDIZED PRIMARY ELECTRON DONOR.

-1- P1M: BINDS FOUR HEME GROUPS PER MOLECULE.

-1- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER

CYTOCHROME C SUBUNITS.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

EMBL: X57597; CAA40820.1; -

DR PIR: S16314; S16314.

DR HSSP: P07173; 6PRC.

DR InterPro: IPR003158; CYTC_RC.

DR InterPro: IPR000345; CYTC_heme_bind.

DR Pfam: PF02276; CytoC_RC; 1.

DR PROSITE: PS00190; CYTOCHROME_C; 2.

KW Electron transport; Photosynthesis; Reaction center; Heme;

KW Duplication. 127 127 HEME 1 (COVALENT).

FT BINDING 130 130 HEME 1 (COVALENT).

FT METAL 131 131 IRON (HEME 1 AXIAL LIGAND).

FT BINDING 178 178 HEME 2 (COVALENT).

FT BINDING 181 181 HEME 1 (COVALENT).

FT METAL 182 182 IRON (HEME 2 AXIAL LIGAND).

FT NON_TER 187 187

SQ SEQUENCE 187 AA; 20681 MW; 8F06B6BEAC557D5 CRC64;

Query Match 79.3%; Score 23; DB 1; Length 187;

Best Local Similarity 83.3%; Pred. No. 48;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
1:111
Db 58 MHVAEF 63

Search completed: October 30, 2002, 12:27:50
Job time : 2.60197 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:30 ; Search time 2.18428 Seconds

(without alignments)
554.401 Million cell updates/sec

Title: US-09-724-571-78

Perfect score: 29

Sequence: 1 VMXVAEF 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	96.6	776	10 Q941L7	Q941L7 populus tre
2	28	96.6	3414	12 Q91G40	Q91G40 langat viru
3	28	96.6	3414	12 Q91G39	Q91G39 langat viru
4	27	93.1	775	10 Q9FE38	Q9FE38 arabidopsi
5	27	93.1	802	10 Q9SUR2	Q9SUR2 arabidopsi
6	26	89.7	324	10 Q9ZP05	Q9ZP05 arabidopsi
7	26	89.7	821	2 Q9XMR4	Q9XMR4 porphyrom
8	25	86.2	201	16 Q34914	Q34914 bacillus su
9	25	86.2	1077	10 P92974	P92974 arabidopsi
10	24	82.8	195	16 Q97F65	Q97F65 clostridium
11	24	82.8	248	5 Q97F65	Q97F65 leishmania
12	24	82.8	249	16 Q9PKD5	Q9PKD5 chlamydia m
13	24	82.8	270	5 Q9NE74	Q9NE74 leishmania
14	24	82.8	317	5 Q9U375	Q9U375 caenorhabdi
15	24	82.8	335	16 Q97RC6	Q97RC6 streptococc
16	24	82.8	367	16 Q97KMO	Q97KMO clostridium

17	24	82.8	378	10 Q9FNG6	Q9FNG6 arabidopsis
18	24	82.8	427	11 Q99L13	Q99L13 mus musculus
19	24	82.8	484	10 Q9C7R2	Q9C7R2 arabidopsis
20	24	82.8	582	17 Q58843	Q58843 pyrococcus
21	24	82.8	584	17 Q9UZN0	Q9UZN0 pyrococcus
22	24	82.8	738	5 P91063	P91063 caenorhabdi
23	24	82.8	817	5 Q9VZJ5	Q9VZJ5 drosophila
24	24	82.8	921	5 Q9K9V0	Q9K9V0 bacillus ha
25	24	82.8	1117	5 Q21477	Q21477 caenorhabdi
26	24	82.8	1122	5 Q61460	Q61460 caenorhabdi
27	24	82.8	1158	5 Q96554	Q96554 caenorhabdi
28	23	79.3	56	13 Q91995	Q91995 xenopus lae
29	23	79.3	57	11 Q9QW56	Q9QW56 mus sp. pro
30	23	79.3	57	13 Q08556	Q08556 phasididae
31	23	79.3	61	4 Q9P166	Q9P166 homo sapien
32	23	79.3	68	13 Q9Y121	Q9Y121 gallus gall
33	23	79.3	89	2 Q92GV2	Q92GV2 yersinia pe
34	23	79.3	105	2 Q91657	Q91657 salmonella
35	23	79.3	129	11 Q63000	Q63000 rattus norv
36	23	79.3	138	13 Q9PSE8	Q9PSE8 gallus gall
37	23	79.3	145	16 Q98BM6	Q98BM6 rhizobium l
38	23	79.3	154	10 Q9FHV3	Q9FHV3 arabidopsis
39	23	79.3	157	6 Q77807	Q77807 bos taurus
40	23	79.3	157	6 Q9NOK5	Q9NOK5 sus scrofa
41	23	79.3	157	6 Q28332	Q28332 callithrix
42	23	79.3	157	6 Q28111	Q28111 bos taurus
43	23	79.3	165	2 Q9F4D2	Q9F4D2 bacteroides
44	23	79.3	172	13 Q91176	Q91176 oryza lat
45	23	79.3	172	13 Q91177	Q91177 oryza lat

ALIGNMENTS

RESULT 1	Q941L7	PRELIMINARY;	PRT;	776 AA.
ID	Q941L7			
AC	Q941L7;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	POTATIVE HIGH-AFFINITY POTASSIUM UPTAKE TRANSPORTER.			
GN	KUP1.			
OS	Populus tremula x Populus tremuloides.			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid 1; Malpighiales; Salicaceae; Populus.			
OX	NCBI_TaxID=47664;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Langer K., Ache P., Fromm J., Hedrich R.;			
RT	"Pitup contributes to wood formation."			
RI	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AJ299422; CAC39168.1; -			
SQ	SEQUENCE 776 AA; 87303 MW; BEC03D57ED0869BC CRC64;			
Query Match	96.6%;	Score 28;	DB 10;	Length 776;
Best Local Similarity	85.7%;	Pred. No. 53;		
Matches	6;	Conservative	0;	Mismatches
			1;	Indels
				Gaps
QY	1 VMXVAEF 7			
DB	604 VMSVAEF 610			
RESULT 2	Q91G40	PRELIMINARY;	PRT;	3414 AA.
ID	Q91G40			
AC	Q91G40;			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	POLYPROTEIN PRECURSOR.			

```

OS  Langat virus (strain TP21).
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC  Flavivirus.
ON  NCBI_TaxID=31638;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  SRRAIN-TP21;
RX  MEDLINE=20192178; PubMed=10725214;
RA  Campbell M.S., Pletnev A.G.;
RT  "Infectious cDNA clones of Langat tick-borne flavivirus that differ
RL  from their parent in peripheral neurovirulence.";
RL  Virology 269:225-237(2000).
DR  EMBL: AF253419; AAF75259.1; -.
DR  HSSP: P14336; 1SVB.
DR  MEROPS: S07.001; -.
DR  InterPro: IPR001410; DEAD.
DR  InterPro: IPR000336; Flavi_glycoprote.
DR  InterPro: IPR001850; Flavi_helicase.
DR  InterPro: IPR001157; Flavi_NSI.
DR  InterPro: IPR001528; Flavi_NS4B.
DR  InterPro: IPR000208; Flavi_NS5.
DR  InterPro: IPR001650; Helicase_C.
DR  Pfam: PF00869; Flavi_glycoprote.1.
DR  Pfam: PF00949; Flavi_helicase.1.
DR  Pfam: PF00949; Flavi_NSI.1.
DR  Pfam: PF01349; Flavi_NS4B.1.
DR  Pfam: PF00972; Flavi_NS5.1.
DR  ProDom: PD001496; Flavi_NSI.1.
DR  ProDom: PD001556; Flavi_glycoprote.1.
DR  SMART: SM00490; HelicC.1.
KW  Signal.
FT  SIGNAL 97 116 POTENTIAL.
FT  CHAIN 1 96 CAPSID PROTEIN.
FT  CHAIN 117 280 MEMBRANE-ASSOCIATED GLYCOPROTEIN
FT  CHAIN PRECURSOR.
FT  CHAIN ENVELOPE MEMBRANE-ASSOCIATED
FT  CHAIN 281 779 GLYCOPROTEIN.
FT  CHAIN 780 1128 NONSTRUCTURAL PROTEIN NS1.
FT  CHAIN 1129 1358 NONSTRUCTURAL PROTEIN NS2A.
FT  CHAIN 1359 1489 NONSTRUCTURAL PROTEIN NS2B.
FT  CHAIN 1490 2110 NONSTRUCTURAL PROTEIN NS3.
FT  CHAIN 2111 2259 NONSTRUCTURAL PROTEIN NS4A.
FT  CHAIN 2260 2511 NONSTRUCTURAL PROTEIN NS4B.
FT  CHAIN 2512 3414 NONSTRUCTURAL PROTEIN NS5.
FT  CHAIN SEQUENCE 3414 AA; 378135 MW; BDE86A46A08BFE73 CRC64;
SQ

Query Match 96.6%; Score 28; DB 12; Length 3414;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   |||||
DB 929 VMTVAEF 935

RESULT 3
Q9IG39 PRELIMINARY; PRT; 3414 AA.
ID Q9IG39;
AC Q9IG39;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE POLYPROTEIN PRECURSOR.
OS Langat virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
ON NCBI_TaxID=11085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATTENUATED STRAIN E5;
RX MEDLINE=20192178; PubMed=10725214;
RA Campbell M.S., Pletnev A.G.;
RT "Infectious cDNA clones of Langat tick-borne flavivirus that differ

```

```

RT  from their parent in peripheral neurovirulence.";
RL  Virology 269:225-237(2000).
DR  EMBL: AF253420; AAF75260.1; -.
DR  HSSP: P14336; 1SVB.
DR  MEROPS: S07.001; -.
DR  InterPro: IPR001410; DEAD.
DR  InterPro: IPR000336; Flavi_glycoprote.
DR  InterPro: IPR001850; Flavi_helicase.
DR  InterPro: IPR001157; Flavi_NSI.
DR  InterPro: IPR001528; Flavi_NS4B.
DR  InterPro: IPR000208; Flavi_NS5.
DR  InterPro: IPR001650; Helicase_C.
DR  Pfam: PF00869; Flavi_glycoprote.1.
DR  Pfam: PF00949; Flavi_helicase.1.
DR  Pfam: PF00949; Flavi_NSI.1.
DR  Pfam: PF01349; Flavi_NS4B.1.
DR  Pfam: PF00972; Flavi_NS5.1.
DR  ProDom: PD001496; Flavi_NSI.1.
DR  ProDom: PD001556; Flavi_glycoprote.1.
DR  SMART: SM00490; HelicC.1.
KW  Signal.
FT  SIGNAL 97 116 POTENTIAL.
FT  CHAIN 1 96 CAPSID PROTEIN.
FT  CHAIN 117 280 MEMBRANE-ASSOCIATED GLYCOPROTEIN
FT  CHAIN PRECURSOR.
FT  CHAIN ENVELOPE MEMBRANE-ASSOCIATED
FT  CHAIN 281 779 GLYCOPROTEIN.
FT  CHAIN 780 1128 NONSTRUCTURAL PROTEIN NS1.
FT  CHAIN 1129 1358 NONSTRUCTURAL PROTEIN NS2A.
FT  CHAIN 1359 1489 NONSTRUCTURAL PROTEIN NS2B.
FT  CHAIN 1490 2110 NONSTRUCTURAL PROTEIN NS3.
FT  CHAIN 2111 2259 NONSTRUCTURAL PROTEIN NS4A.
FT  CHAIN 2260 2511 NONSTRUCTURAL PROTEIN NS4B.
FT  CHAIN 2512 3414 NONSTRUCTURAL PROTEIN NS5.
FT  CHAIN SEQUENCE 3414 AA; 378014 MW; BF8F8ACEB96D534 CRC64;
SQ

Query Match 96.6%; Score 28; DB 12; Length 3414;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
   |||||
DB 929 VMTVAEF 935

RESULT 4
Q9FE38 PRELIMINARY; PRT; 775 AA.
ID Q9FE38;
AC Q9FE38;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TINY ROOT HAIR 1 PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
ON NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rigas S., Debrosses G., Haralampidis K., Vicente-Aguillo F.,
RA Feldmann K.A., Grabov A., Dolan L., Hatzopoulos P.;
RT "A Potassium Transporter Required for Tip Growth in Arabidopsis.";
RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
RX EMBL: AJ296156; CAC16138.1; -.
RX EMBL: AJ296155; CAC16137.1; -.
DR InterPro: IPR003855; K+-transprt.
DR Pfam: PF02705; K.trans.1.
DR SEQUENCE 775 AA; 86842 MW; B0C5068B48E8180 CRC64;
SQ

Query Match 93.1%; Score 27; DB 10; Length 775;
Best Local Similarity 71.4%; Pred. No. 99;

```

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
11:1111
DB 602 VMSIAEF 608

RESULT 5

Q9SUR2 PRELIMINARY; PRT; 802 AA.

AC Q9SUR2;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE PUTATIVE POTASSIUM TRANSPORT PROTEIN.
GN F9D16.110 OR A74G23640.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;

RN [1]
RA SEQUENCE FROM N.A.
RA Bevan M., Medler H., Medler E., Wambutt R., Hohnsels J., Mewes H.W.,
RA Mayer K.F.X., Scheller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RA SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RA SEQUENCE FROM N.A.
RA Medler H., Medler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RA SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL055394; CAA23030.1; -
DR EMBL: AL161559; CAB79319.1; -
DR InterPro: IPR003855; K+-transprt.
DR Pfam: PF02705; K_trans.1.
SQ SEQUENCE 802 AA; 89817 MW; 617AF5F76B99B60 CRC64;

Query Match 93.1%; Score 27; DB 10; Length 802;
Best Local Similarity 71.4%; Pred. No. 1e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
11:1111
DB 629 VMSIAEF 635

RESULT 6

Q92PQ5 PRELIMINARY; PRT; 324 AA.

AC Q92PQ5;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE PUTATIVE STEROID SULFOTRANSFERASE.
GN AT2G03770.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;

RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN=CV, COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,

RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Talton L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhagen G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana";
RL Nature 402:761-768(1999).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006836; AAD20079.1; -
DR HSSP: P50224; 1C3M.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer; 1.
DR Prodom: PD001218; Sulfotransfer; 1.
KW Transferase.
SQ SEQUENCE 324 AA; 37717 MW; B486C799AB13E1D3 CRC64;

Query Match 89.7%; Score 26; DB 10; Length 324;
Best Local Similarity 71.4%; Pred. No. 78;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
11:1111
DB 233 VMRIAEF 239

RESULT 7

Q9XBW4 PRELIMINARY; PRT; 821 AA.

AC Q9XBW4;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE IMMUNOREACTIVE 92 KDA ANTIGEN Pg21.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
NCBI_TaxID=837;
RN [1]
RA SEQUENCE FROM N.A.
RA Ross B.C., Barr I., Patterson M., Agius C., Rothel L., Margetts M.,
RA Hocking D., Webb E.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF153768; AAD38980.1; -
DR MEROPS: M14.0PB; -
DR InterPro: IPR000601; PKD_domain.
DR Pfam: PF00801; PKD; 3.
DR Pfam: PF00246; Zn_carpodpept; 1.
DR PRINTS: PR00765; CRBOXYPTASEA.
DR SMART: SM00089; PKD; 3.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
DR PROSITE: PSS0093; PKD; 3.
SQ SEQUENCE 821 AA; 91517 MW; 250843B2C9C83E2 CRC64;

Query Match 89.7%; Score 26; DB 2; Length 821;
Best Local Similarity 85.7%; Pred. No. 2e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
11:1111
DB 565 VMPVAEF 571

RESULT 8

Q34914 PRELIMINARY; PRT; 201 AA.

AC 034914;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, last annotation update)
 DE YTBQ.
 GN Bacillus subtilis.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9631235; PubMed=8763940;
 RA Bower S., Perkins J.B., Yocum R.R., Howitt C.L., Rahaim P., Pero J.,
 RT "Cloning, sequencing, and characterization of the Bacillus subtilis
 RT biotin biosynthetic operon."
 RL J. Bacteriol. 178:4122-4130(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich D.,
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=9804403; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertorello M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano N.J., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denton F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Enliant K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goftau A., Gollightly E.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Kluett-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogilawa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serrier P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Mamuit R., Wedler E., Wedler H., Weltzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 RT subtilis."
 RL Nature 390:249-256(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF008220; AAC00267.1; -
 DR EMBL: Z99119; CAB14996.1; -
 KW Complete proteome.
 SQ SEQUENCE 201 AA; 22733 MW; CBFAD9B6BA348F06 CRC64;

Query Match 86.2%; Score 25; DB 16; Length 201;
 Best Local Similarity 57.1%; Pred. No. 90;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXYAEF 7
 : : : : :
 DB 23 IMDIAEF 29

RESULT 9
 ID P92974 PRELIMINARY; PRT; 1077 AA.
 AC P92974;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
 DE UBIOUITIN-ACTIVATING ENZYME E1.
 GN UBA1 OR UBA2.
 OS Arabidopsis thaliana (Mouse-ear cress). Embryophyta; Tracheophyta;
 OC Eukaryota; Viridiplantae; Streptophyta; Magnoliopsida; core eudicots; Rosidae;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ECOTYPE COLUMBIA; T.C., Vierstra R.D.;
 RA Hatfield P.M., Carpenter
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=9806901; PubMed=9405937;
 RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
 RT Sequence features of the regions of 1,044,062 bp covered by thirteen
 RT physically assigned pl clones."
 RL DNA Res. 4:291-300(1997)
 CC -1 FUNCTION: ACTIVATES UBIOUITIN BY FIRST ADENYLATING WITH ATP ITS
 CC CARBOXY-TERMINAL GLYCINE RESIDUE AND THEREAFTER LINKING THIS
 CC RESIDUE TO THE SIDE CHAIN OF A CYSTEINE RESIDUE IN E1, YIELDING AN
 CC UBIOUITIN-E1 THIOLESTER AND FREE AMP (BY SIMILARITY).
 CC -1 PATHWAY: FIRST STEP IN UBIOUITIN CONJUGATION.
 CC -1 SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1 SIMILARITY: TO E1 IN OTHER SPECIES.
 DR EMBL: U40566; AB37569.1; -
 DR EMBL: AB006700; BAB08968.1; -
 DR InterPro: IPR000734; Lipase.
 DR InterPro: IPR000205; NAD_binding.
 DR InterPro: IPR000594; Thf_family.
 DR InterPro: IPR000127; Ubact_repeat.
 DR InterPro: IPR000011; Ubiquitin_activ_enz.
 DR Pfam: PF00899; Thf_family_2.
 DR Pfam: PF02134; Ubact_repeat_1.
 DR PROSITE: PS00120; LIPASE_SER; UNKNOWN_1.
 DR PROSITE: PS00536; UBIOUITIN_ACTIVAT_1; 1.
 DR PROSITE: PS00865; UBIOUITIN_ACTIVAT_2; 1.
 KW ligase; Ubiquitin conjugation.
 FT ACT_SITE 653
 SQ SEQUENCE 1077 AA; 119622 MW; CE3A36A9A9A218 CRC64;

Query Match 86.2%; Score 25; DB 10; Length 1077;
 Best Local Similarity 71.4%; Pred. No. 4.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXYAEF 7
 : : : : :
 DB 54 VMSMAEF 60

RESULT 10
 ID Q97F65 PRELIMINARY; PRT; 195 AA.
 AC Q97F65;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
 DE THYMIDINE KINASE.
 GN CAC2887.
 OS Clostridium acetobutylicum.

OC Bacteria: Firmicutes: Bacillus/Clostridium group: Clostridiaceae;
 OC Clostridium;
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum";
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL: AE007786; AAK80830.1; -;
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR001267; Thymd_kin_cell.
 DR Pfam: PF00265; TK; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS00603; TK_CELULAR_TYPE; UNKNOWN_1.
 KM Kinase: Complete proteome.
 SQ SEQUENCE 195 AA; 22303 MW; 6EDB51986B203BA CRC64;

Query Match
 Best Local Similarity 57.1%; Score 24; DB 16; Length 195;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 : | | | |
 Db 131 LMAIAEF 137

RESULT 11
 ID 097008 PRELIMINARY; PRT; 248 AA.
 AC 097008;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 26.4 KDA PROTEIN.
 GN L7610.05.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
 RA Ralendream M.A., Barrell B.G.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RX MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome.";
 RL Genome Res. 8:135-145(1998).
 DR EMBL: AL139794; CAC22635.1; -;
 KM Hypothetical protein.
 SQ SEQUENCE 248 AA; 26365 MW; A80440E18B3F69B5 CRC64;

Query Match
 Best Local Similarity 82.8%; Score 24; DB 5; Length 248;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 : | | | |
 Db 196 LMRVAEF 202

RESULT 12
 O9PKDS

ID 09PKDS PRELIMINARY; PRT; 249 AA.
 AC 09PKDS;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SERINE/THREONINE PROTEIN PHOSPHATASE, PUTATIVE.
 GN TC0530.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MORN / NIGG;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uitterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Morn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 DR EMBL: AE002321; AAF39372.1; -;
 DR TIGR: TC0530; -;
 DR InterPro: IPR001933; PP2C_domain.
 DR InterPro: IPR003643; Putat_PPase.
 DR ProDom: PD006823; Putat_PPase.
 DR SMART: SM00352; PP2C; 1.
 DR SMART: SM00331; PP2C_SIG; 1.
 KM Complete proteome.
 SQ SEQUENCE 249 AA; 28192 MW; 6BD227BD3281164F CRC64;

Query Match
 Best Local Similarity 82.8%; Score 24; DB 16; Length 249;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VMXVAEF 7
 : | | | |
 Db 1 MMRVAEF 7

RESULT 13
 ID 09NE74 PRELIMINARY; PRT; 270 AA.
 AC 09NE74;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 29.3 KDA PROTEIN.
 GN L5075.02.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RA Beck A., Klages S., Reinhardt R., Ivens A.C., Quail M.,
 RA Ralendream M.A., Barrell B.G.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RX MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome.";
 RL Genome Res. 8:135-145(1998).
 DR EMBL: AL163552; CAB86964.1; -;
 KM Hypothetical protein.
 SQ SEQUENCE 270 AA; 29335 MW; 5B91DA37A16FC3CB CRC64;

Query Match
 Best Local Similarity 82.8%; Score 24; DB 5; Length 270;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
11:1111

Db 148 VMLAEF 154

RESULT 14

Q9U375 PRELIMINARY; PRT; 317 AA.
AC Q9U375:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE T08D2.2 PROTEIN.
GN T08D2.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurtry A.A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z92839; CAB07417.2; -.
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
SQ SEQUENCE 317 AA; 36591 MW; 145E8B02B75AE42 CRC64;

Query Match 82.8%; Score 24; DB 5; Length 317;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MXVAEF 7
111111

Db 223 MSVAEF 228

RESULT 15

Q97RC6 PRELIMINARY; PRT; 335 AA.
AC Q97RC6:
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE 6-PHOSPHOFRUCTOKINASE.
GN SP0896.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIGR4:
MEDLINE=21357209; PubMed=11463916;
RA Tettein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D., C.L.,
RA Holtzapple E., Khouli H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblum T.V., Angiolli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
pneumoniae.";
RL Science 293:498-506(2001).

DR EMBL; AE007395; AAK75023.1; -.
DR TIGR; SP0896; -.
DR InterPro: IPR000023; Phosphofructokinase.
DR Pfam: PF00365; PFK; 1.
DR PRINTS; PR00476; PHERCTKINASE.
DR Prodom: PD000707; Phosphofructokinase; 1.
DR PROSITE: PS00433; PHOSPHOFRUCTOKINASE; 1.
KW Kinase; Complete proteome.
SQ SEQUENCE 335 AA; 35174 MW; 2DFBAAC58CAF66 CRC64;

Query Match 82.8%; Score 24; DB 16; Length 335;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VMXVAEF 7
111111

Db 224 VMSAEF 230

Search completed: October 30, 2002, 12:30:07
Job time : 6.18428 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:30 ; Search time 2.80835 Seconds
(without alignments)
554.401 Million cell updates/sec

Title: US-09-724-571-91

Perfect score: 42

Sequence: 1 SEVNLIAR 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

1: SP.REMBL.19:*
2: SP.bacteria:*
3: SP.fungi:*
4: SP.human:*
5: SP.invertebrate:*
6: SP.mammal:*
7: SP.mhc:*
8: SP.organelle:*
9: SP.phage:*
10: SP.plant:*
11: SP.todent:*
12: SP.virus:*
13: SP.vertebrate:*
14: SP.unclassified:*
15: SP.virus:*
16: SP.bacteriap:*
17: SP.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	83.3	649	5	Q9V7X7 drosophila
2	35	83.3	948	12	Q9QW73
3	34	81.0	529	12	Q39281 avian adeno
4	33	78.6	668	16	Q97KW1 equine hepr
5	33	78.6	837	4	Q9UH51 clostridium
6	32	76.2	132	16	Q9WZB6 homo sapien
7	32	76.2	279	10	Q65159 thermotoga
8	32	76.2	422	10	Q65159 phaseolus v
9	32	76.2	467	10	Q49336 spinacia ol
10	32	76.2	467	10	Q9SEB8 phaseolus a
11	32	76.2	472	10	Q41688 vigna radia
12	32	76.2	472	10	Q24544 phaseolus a
13	32	76.2	476	16	Q41687 vigna radia
14	32	76.2	942	10	Q92512 rhizobium m
15	31	73.8	151	10	Q96585 avian adeno
16	31	73.8	167	2	Q9FL93 arabisdopsis
					P77934 photobacter

17	31	73.8	175	10	Q93VA3	Q93va3 arabisdopsis
18	31	73.8	210	10	Q9FKX3	Q9fkx3 arabisdopsis
19	31	73.8	225	12	Q84614	Q84614 parametium
20	31	73.8	246	13	Q9W6E5	Q9w6e5 gallus gall
21	31	73.8	295	10	Q9SHP7	Q9shp7 arabisdopsis
22	31	73.8	384	4	Q9GQNS	Q9gqns homo sapien
23	31	73.8	389	13	Q9W6Z3	Q9w6z3 xenopus lae
24	31	73.8	393	11	Q91XQ1	Q91xq1 rattus norv
25	31	73.8	394	16	Q9A4F6	Q9a4f6 caulobacter
26	31	73.8	397	2	Q9S523	Q9s523 escherichia
27	31	73.8	411	2	Q68771	Q68771 yersinia pe
28	31	73.8	411	2	Q935E6	Q935e6 salmonella
29	31	73.8	435	3	Q74765	Q74763 schizosacch
30	31	73.8	442	13	Q9W6Z2	Q9w6z2 xenopus lae
31	31	73.8	443	11	Q921N9	Q921n9 mus musculu
32	31	73.8	443	11	Q91Z18	Q91z18 mus musculu
33	31	73.8	906	5	Q17281	Q17281 botryllus s
34	31	73.8	1357	12	Q89328	Q89328 rice ragged
35	31	73.8	2270	12	Q9JFN3	Q9jfn3 tupia para
36	31	73.8	3169	2	Q9JFN3	Q9jfn3 escherichia
37	31	73.8	3169	2	Q9ZGR4	Q9zgr4 escherichia
38	30.5	72.6	557	13	Q90Z52	Q90z52 brachydanio
39	30	71.4	118	16	Q988K6	Q988k6 rhizobium l
40	30	71.4	133	16	Q9CRU2	Q9cru2 lactococcus
41	30	71.4	218	8	Q98455	Q98455 spirogyra m
42	30	71.4	293	5	Q9VJE7	Q9vje7 drosophila
43	30	71.4	298	11	Q9EQ06	Q9eq06 mus musculu
44	30	71.4	300	4	Q9UKD4	Q9ukd4 homo sapien
45	30	71.4	300	4	Q96HP6	Q96ht6 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9V7X7	PRELIMINARY:	PRT:	649 AA.
AC	Q9V7X7:			
DT	01-MAY-2000 (TREMblrel. 13, Created)			
DT	01-MAY-2000 (TREMblrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMblrel. 19, Last annotation update)			
DE	CG8961 PROTEIN (LD32631P).			
OS	TEF OR CG8961.			
GN	Drosophila melanogaster (Fruit fly)			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY.			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Ceoliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,			
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	April J.F., Agbayani A., An H.-O., Andrews-Plankk C., Baldwin D.,			
RA	Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,			
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin M., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,			

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Pollard J., Puri V., Reese M.G.,
 RA Palazzolo M., Plattman K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Reinert K., Remington K., Simpson M., Skupski M.P., Smith T.,
 RA Shue B.C., Siden-Kiamos I., Stapleton M., Strong R., Sun E.,
 RA Spier E., Spradling A.C., Stapleton M., Venter E., Wang A.H., Wang X.,
 RA Svirskas R., Turner C., Turner R., Venter E., Weissbach J.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu X.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y, CN BW SP;
 RA Stapleton M., Brokstein P., Hong L., Abhayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacle J., Rubin G.M., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RA Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003804; AAF57911.1; -;
 DR EMBL; AY051818; AAK93242.1; -;
 DR Flybase; FBgn0043472; tef.
 DR InterPro; IPR000822; znf-C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR SMART; SM00355; znf-C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
 DR DNA-binding; Metal-binding; Zinc-finger.
 KW SEQUENCE 649 AA; 73357 MW; F96518EB37D60E68 CRC64;
 SQ

Query Match 83.3%; Score 35; DB 5; Length 649;
 Best Local Similarity 87.5%; Pred. No. 27;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SEVNLAE 8
 ||:||||
 Db 131 SEINLAE 138

RESULT 2
 O9QM73 PRELIMINARY; PRT; 948 AA.
 AC O9QM73;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE P1.
 OS Avian adenovirus type 8 (strain ATCC A-2a) (Fowl adenovirus 8).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
 OX NCBI_TaxID=66295;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20318671; PubMed=10859390;
 RX Ojikic D., Nagy E.;
 RT "The complete nucleotide sequence of fowl adenovirus type 8.";
 RL J. Gen. Virol. 81:1833-1837(2000).
 [2]
 RN SEQUENCE FROM N.A.
 RP Ojikic D., Nagy E.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF083975; AAD50344.2; -;
 DR HSSP; P03277; IDHX.
 DR InterPro; IPR000736; Adeno_hexon.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF01065; Adeno_hexon; 1.
 DR ProDom; PD002815; Adeno_hexon; 1.

DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 SQ SEQUENCE 948 AA; 107604 MW; 67146E79F31446B8 CRC64;

Query Match 83.3%; Score 35; DB 12; Length 948;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAE 9
 |||||:
 Db 587 SEVNLAE 595

RESULT 3
 O39281 PRELIMINARY; PRT; 529 AA.
 ID O39281;
 AC O39281;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE COUNTERPART OF HSV-1 GENE UL21 AND VZV GENE 38.
 GN 40.
 OS Equine herpesvirus 4.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10331;
 RN [1]
 RP SEQUENCE OF 489-529 FROM N.A.
 RC STRAIN=NS80567;
 RX MEDLINE=90362066; PubMed=2167933;
 RA Nicolson L., Cullinane A.A., Onions D.E.;
 RT "The nucleotide sequence of an equine herpesvirus 4 gene homologue of
 the herpes simplex virus 1 glycoprotein H gene.";
 RL J. Gen. Virol. 71:1793-1800(1990).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=NS80567;
 RC MEDLINE=98264497; PubMed=9603335;
 RX Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-4.";
 RL J. Gen. Virol. 79:1197-1203(1998).
 [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=NS80567;
 RC Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF030027; AAC59558.1; -;
 KW SEQUENCE 529 AA; 58150 MW; E2B737B89D489413 CRC64;
 SQ

Query Match 81.0%; Score 34; DB 12; Length 529;
 Best Local Similarity 87.5%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 EVNLAE 9
 | |||||
 Db 423 EVNLAE 430

RESULT 4
 O97KW1 PRELIMINARY; PRT; 668 AA.
 ID O97KW1;
 AC O97KW1;
 DT 01-OCT-2001 (TReMBLrel. 18, Created)
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE METHYL-ACCEPTING CHEMOTAXIS PROTEIN.
 GN CAC0805.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium *Clostridium acetobutylicum*.";
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL, AE007596; AAK78781.1; -.
 DR InterPro: IPR004010; Cache.
 DR InterPro: IPR004089; Chemotaxis_transducer.
 DR InterPro: IPR003660; HAMP.
 DR InterPro: IPR003122; TsrH.
 DR Pfam: PF02743; Cache; 1.
 DR Pfam: PF00672; HAMP; 1.
 DR Pfam: PF00015; MCPsignal; 1.
 DR SMART: SM00304; HAMP; 1.
 DR SMART: SM00283; MA; 1.
 DR SMART: SM00319; TarH; 1.
 KM Complete proteome.
 SO SEQUENCE 668 AA; 73157 MW; F897FD68E105D6FF CRC64;

Query Match 78.6%; Score 33; DB 16; Length 668;
 Best Local Similarity 75.0%; Pred. No. 80;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SEVNLAE 8
 DB 71 SEINLAE 78

RESULT 5
 ID O9UH51 PRELIMINARY; PRT; 837 AA.
 AC O9UH51;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 13, Last annotation update)
 DE LOW DENSITY LIPOPROTEIN RECEPTOR (FRAGMENT).
 GN LDLR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA Van Leuven F., Thiry E., Stas L., Lambrechts M., Boon T.,
 RA Bruynseels K., Descamps O.;
 RT "Sequencing the coding exons for mutational analysis of the LRP1 and
 RT LDLR genes on individual DNA samples.";
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL, AF217403; AAF24515.1; -.
 DR HSSP: P01130; LDL.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000581; EGF-like.
 DR InterPro: IPR001881; EGF-like.
 DR InterPro: IPR000033; Ldl_receptor_rep.
 DR InterPro: IPR002172; Ldl_recept_A.
 DR Pfam: PF00008; EGF; 3.
 DR Pfam: PF00057; Ldl_recept_a; 7.
 DR Pfam: PF00058; Ldl_recept_b; 5.
 DR PRINTS: PR00261; LDLRECEPTOR.
 DR SMART: SM00179; EGF_CA; 1.
 DR SMART: SM00001; EGF_like; 2.
 DR SMART: SM00192; LDla; 7.
 DR SMART: SM00135; LY; 5.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR PROSITE: PS01209; LDLRA_1; 7.
 DR PROSITE: PS00068; LDLRA_2; 7.
 KM Calcium-binding; EGF-like domain; Glycoprotein; Lipoprotein; Receptor;
 Repeat.

FT NON_TER 1 1
 SQ SEQUENCE 837 AA; 92914 MW; 22AFB9382D7CE656 CRC64;

Query Match 78.6%; Score 33; DB 4; Length 837;
 Best Local Similarity 87.5%; Pred. No. 1e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SEVNLAE 8
 DB 614 SDVNLAE 621

RESULT 6
 ID O9WZB6 PRELIMINARY; PRT; 132 AA.
 AC O9WZB6;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN TM0648.
 GN Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smit H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of *Thermotoga maritima*.";
 RT Nature 399:323-329(1999).
 DR EMBL, AE001738; AAD35732.1; -.
 DR TIGR: TM0648; -.
 DR InterPro: IPR002934; NTP_transf.
 DR Pfam: PF01909; NTP_Transf_2; 1.
 KM Complete proteome.
 SO SEQUENCE 132 AA; 15264 MW; 4E973A6234276462 CRC64;

Query Match 76.2%; Score 32; DB 16; Length 132;
 Best Local Similarity 66.7%; Pred. No. 25;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAE 9
 DB 86 SEINLAE 94

RESULT 7
 ID O65159 PRELIMINARY; PRT; 279 AA.
 AC O65159;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 1-AMINOACLOPROPANE-1-CARBOXYLIC ACID SYNTHASE (FRAGMENT).
 GN ACS1.
 OS Phaseolus vulgaris (Kidney bean) (French bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
 OX NCBI_TaxID=3885;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. TAYLOR;
 RA Pidgeon C.M., Reid D.M.;
 RT "The gene of light in controlling ACC synthase activity and gene
 expression in *Phaseolus vulgaris* cv. Taylor.";

Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC ONE AND A MITOCHONDRIAL ONE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.
 CC EMBL: AF053355; AAC08356.1; -.
 DR HSSP: P37821; 1B8G.
 DR InterPro: IPR001176; ACC_synthase.
 DR InterPro: IPR001511; Aminotran_1.
 DR Pfam: PF00155; aminotran_1_2; 1.
 DR PRINTS: PR00753; ACCSYNTHASE.
 DR PROSITE: PS00105; AA_TRANSFER_CLASS_1; 1.
 DR Pyridoxal phosphate. 1
 FT NON_TER 1 279
 FT SEQUENCE 279 AA; 31069 MW; 50AE1FF551CFA7 CRC64;
 SQ
 Query Match 76.2%; Score 32; DB 10; Length 279;
 Best Local Similarity 66.7%; Pred. No. 54;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAE 9
 DB 129 SEVNLAE 137

RESULT 8
 OY 049936 PRELIMINARY; PRT; 422 AA.
 AC 049936;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE BASIC LEUCINE ZIPPER PROTEIN.
 GN BZIP.
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OC NCBI_TaxID=3562;
 OX [1]
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bolle C., Lieberstedt T., Herranen M., Herrmann R., Oelmueller R.;
 RT "Molecular characterization of the spinach G-box binding protein family.";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY.
 CC EMBL: AJ223624; CA011499.1; -.
 DR HSSP: P03069; 1DGC.
 DR InterPro: IPR001871; bZIP.
 DR Pfam: PF00170; bZIP_1.
 DR SMART: SM00338; BZIP_1.
 DR PROSITE: PS00036; BZIP_BASIC; 1.
 DR DNA-binding; Nuclear protein.
 KW SEQUENCE 422 AA; 44924 MW; F9713C3BCA65D2F0 CRC64;
 SQ
 Query Match 76.2%; Score 32; DB 10; Length 422;
 Best Local Similarity 87.5%; Pred. No. 83;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAE 8
 DB 324 SEVNLAE 331

RESULT 9
 OY 09SEB8 PRELIMINARY; PRT; 467 AA.
 ID 09SEB8;
 AC 09SEB8;
 DT 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE 1-AMINOCYCLOPROPANE-1-CARBOXYLIC ACID SYNTHASE.
 GN ACS7.
 OS Phaseolus aureus (Mung bean) (Vigna radiata).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 OC NCBI_TaxID=3916;
 OX [1]
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-20074367; PubMed-10608655;
 RX Yi H.C., Joo S., Nam K.H., Lee J.S., Kang B.G., Kim W.T.;
 RT "Auxin and brassinosteroid differentially regulate the expression of three members of the 1-aminocyclopropene-1-carboxylate synthase gene family in mung bean (Vigna radiata L.).";
 RT Plant Mol. Biol. 41:443-454(1999).
 RL
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC ONE AND A MITOCHONDRIAL ONE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.
 CC EMBL: AF151961; A041083.1; -.
 DR HSSP: P37821; 1B8G.
 DR InterPro: IPR001176; ACC_synthase.
 DR InterPro: IPR001511; Aminotran_1.
 DR InterPro: IPR000515; BPD_transp.
 DR Pfam: PF00155; aminotran_1_2; 1.
 DR PRINTS: PR00753; ACCSYNTHASE.
 DR PROSITE: PS00105; AA_TRANSFER_CLASS_1; 1.
 DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
 DR Pyridoxal phosphate. 1
 KW SEQUENCE 467 AA; 53044 MW; 121423C4ABACB079 CRC64;
 SQ
 Query Match 76.2%; Score 32; DB 10; Length 467;
 Best Local Similarity 66.7%; Pred. No. 93;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAE 9
 DB 211 SEVNLAE 219

RESULT 10
 OY 041688 PRELIMINARY; PRT; 467 AA.
 ID 041688;
 AC 041688;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE 1-AMINOCYCLOPROPANE-1-CARBOXYLIC ACID SYNTHASE.
 GN VRACS7.
 OS Vigna radiata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 OC NCBI_TaxID=157791;
 OX [1]
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-ETIOLOGATED HYPOCOTYL;
 RA Kim W.;
 RT Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC ONE AND A MITOCHONDRIAL ONE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.
 CC EMBL: U34987; AA78274.1; -.
 DR HSSP: P37821; 1B8G.
 DR InterPro: IPR001176; ACC_synthase.
 DR InterPro: IPR001511; Aminotran_1.

DR InterPro: IPR000515; BPD_transp.
 DR Pfam: PF00155; aminotran_1.2; 1.
 DR PRINTS: PR00753; ACCSYNTHASE.
 DR PROSITE: PS00105; AA_TRANSFER_CLASS_1; 1.
 DR PROSITE: PS00402; BP_TRANSF_INN_MEMBER; UNKNOWN_1.
 KW Pyridoxal phosphate. 467 AA; 53058 MW; 16710824A0376A79 CRC64;
 SO SEQUENCE

Query Match 76.2%; Score 32; DB 10; Length 467;
 Best Local Similarity 66.7%; Pred. No. 93;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
 ||:||||:|
 Db 211 SELNLVDF 219

RESULT 11

024544 PRELIMINARY; PRT; 472 AA.
 AC 024544;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE..
 GN VR-ACS6.
 OS Phaseolus aureus (Mung bean) (Vigna radiata).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 OX NCBI_TaxID=3916;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ETIOLATED HYPOCOTYL;
 RX MEDLINE=97294927; PubMed=9150600;
 RA Yoon I., Mori H., Kim J., Kang B., Imaseki H.;
 RT "VR-ACS6 is an auxin-inducible 1-aminocyclopropene-1-carboxylate
 RT synthase gene in mungbean (Vigna radiata).";
 RL Plant Cell Physiol. 38:217-224(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HYPOCOTYL;
 RA Yoon I.S., Park D.H., Mori H., Imaseki H., Kang B.G.;
 RT "Characterization of an auxin-inducible ACC synthase gene, VR-ACS6 in
 RT mungbean and expression of its promoter in transgenic tobacco.";
 RL Submitted (Oct-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- CORPACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
 CC ONE AND A MITOCHONDRIAL ONE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC EMBL: AB000679; BAA19161.1; -;
 DR EMBL: AB018355; BAA33859.1; -;
 DR HSSP: P37821; 188G.
 DR InterPro: IPR001176; ACC_synthase.
 DR InterPro: IPR001511; Aminotran_1.
 DR Pfam: PF00155; aminotran_1.2; 1.
 DR PRINTS: PR00753; ACCSYNTHASE.
 DR PROSITE: PS00105; AA_TRANSFER_CLASS_1; 1.
 KW Pyridoxal phosphate.
 SO SEQUENCE 472 AA; 53572 MW; 1404AD4B3BD0421 CRC64;

Query Match 76.2%; Score 32; DB 10; Length 472;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
 ||:||||:|
 Db 211 SELNLVDF 219

RESULT 12

041687 PRELIMINARY; PRT; 472 AA.
 AC 041687;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 1-AMINOCYCLOPROPANE-1-CARBOXYLIC ACID SYNTHASE.
 GN VRACS6.
 OS Vigna radiata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 OX NCBI_TaxID=157791;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ETIOLATED HYPOCOTYLS;
 RA Kim W.;
 RL Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
 CC ONE AND A MITOCHONDRIAL ONE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC EMBL: U34986; AAA78273.1; -;
 DR HSSP: P37821; 188G.
 DR InterPro: IPR001176; ACC_synthase.
 DR InterPro: IPR001511; Aminotran_1.
 DR Pfam: PF00155; aminotran_1.2; 1.
 DR PRINTS: PR00753; ACCSYNTHASE.
 DR PROSITE: PS00105; AA_TRANSFER_CLASS_1; 1.
 KW Pyridoxal phosphate.
 SO SEQUENCE 472 AA; 53518 MW; 0C5F3E22AB0E8EAF CRC64;

Query Match 76.2%; Score 32; DB 10; Length 472;
 Best Local Similarity 66.7%; Pred. No. 94;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
 ||:||||:|
 Db 211 SELNLVDF 219

RESULT 13

092ZL2 PRELIMINARY; PRT; 476 AA.
 AC 092ZL2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE NIFE OXIDOREDUCTASE.
 GN NIFE.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC plasmid pSyma (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN=1021;
 RC MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barlow-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
 RA Gurjal M., Hong A., Huzar L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Kalmann S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Voh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti pSyma megaplasmid.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 DR EMBL: AE007236; AAK65110.1; -;
 KW Plasmid; Complete proteome.
 SO SEQUENCE 476 AA; 52358 MW; 83ECE94BA814C713 CRC64;

Query Match 76.2%; Score 32; DB 16; Length 476;
 Best Local Similarity 55.6%; Pred. No. 95;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAEF 9
 ID 195 SDINIGEF 203

RESULT 14

Q96585 PRELIMINARY; PRT; 942 AA.
 AC Q96585;
 DT 01-FEB-1997 (TREMblrel. 02, Created)
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE VIRAL STRUCTURAL PROTEIN.
 GN HEXON.
 OS Avian adenovirus gall (strain Phelps) (Powl adenovirus 1) (CELO).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Adenovirus.
 OX NCBI_TaxID=10553;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FAV1 (CELO);
 RX MEDLINE=97049067; PubMed=8893797;
 RA Akopian T.A., Doronin K.K., Karpov V.A., Naroditsky B.S.;
 RT "Sequence of the avian adenovirus FAV1 (CELO) DNA encoding the hexon-associated protein pIV and hexon.";
 RL Arch. Virol. 141:1759-1765(1996).
 DR EMBL; Z67970; CAA91908.1; -.
 DR HSP; P03277; IDHX.
 DR InterPro; IPR00736; Adeno_hexon.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF01065; Adeno_hexon; 1.
 DR ProDom; PD002815; Adeno_hexon; 1.
 DR PROSITE; PS00678; WD_REPEATS.1; UNKNOWN.1.
 SQ SEQUENCE 942 AA; 10665 MW; 903732D3F5E6DC6F CRC64;

Query Match 76.2%; Score 32; DB 12; Length 942;
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAEF 9
 ID 581 TEVNLMAEF 589

RESULT 15

Q9FL93 PRELIMINARY; PRT; 151 AA.
 AC Q9FL93;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE GENOMIC DNA, CHROMOSOME 5, TAC CLONE:K21C13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA.
 RX MEDLINE=98344145; PubMed=9679202;
 RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence features of the regions of 1,381,565 bp covered by twenty one physically assigned pl and TAC clones.";
 RL DNA Res. 5:131-145(1998).
 DR EMBL; AB010693; BAB10867.1; -.
 DR HSP; P08197; ICYJ.
 DR InterPro; IPR000345; CytC_heme_bind.

DR InterPro; IPR003088; Cyt_C1.
 DR Pfam; PF00034; CYTOCHROME_C; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN.1.
 SQ SEQUENCE 151 AA; 17031 MW; 2C7F3A4E4ADEDE2 CRC64;

Query Match 73.8%; Score 31; DB 10; Length 151;
 Best Local Similarity 75.0%; Pred. No. 48;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVNLAEF 9
 ID 129 EIKLAEF 136

Search completed: October 30, 2002, 12:30:13
 Job time : 5.80835 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:10 ; Search time 0.773956 Seconds
(without alignments)
450.253 Million cell updates/sec

Title: US-09-724-571-91
Perfect score: 42
Sequence: 1 SEVNILAEF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	78.6	860	1	LDLR_HUMAN
2	32	76.2	476	1	NIFE_RHIME
3	32	76.2	942	1	HEX_ADECI
4	31	73.8	181	1	ORX_VIBCH
5	31	73.8	445	1	NEK2_MOUSE
6	31	73.8	445	1	NEK2_HUMAN
7	30	71.4	103	1	Y041_AQUAE
8	30	71.4	328	1	Y778_AQUAE
9	30	71.4	581	1	Y015_SCHPO
10	30	71.4	768	1	MEC4_CAEEL
11	30	71.4	769	1	MEC4_CAEEL
12	30	71.4	854	1	LDLR_CRIGR
13	30	71.4	864	1	LDLR_MOUSE
14	30	71.4	890	1	IF2_CHLPN
15	29	69.0	135	1	Y549_RICPR
16	29	69.0	230	1	RR3_PORPU
17	29	69.0	338	1	YABT_BACSU
18	29	69.0	405	1	CPXK_SACER
19	29	69.0	460	1	UDPE_GVLO
20	29	69.0	613	1	ENP4_MOUSE
21	29	69.0	616	1	ENP4_HUMAN
22	29	69.0	1666	1	CO3_CAVPO
23	29	69.0	5430	1	ACE7_HUMAN
24	28	66.7	237	1	ORN_HUMAN
25	28	66.7	250	1	Y006_ANASP
26	28	66.7	302	1	BIRA_HAEIN
27	28	66.7	306	1	BUB2_YEAST
28	28	66.7	338	1	CYSP_PCOIT
29	28	66.7	338	1	CYSP_PCOIT
30	28	66.7	338	1	CYSP_PCOIT
31	28	66.7	391	1	Y773_CAEEL
32	28	66.7	394	1	ACKA_LACSK
33	28	66.7	419	1	P47K_PSECL

34	28	66.7	437	1	V014_FOWPY	Q9J517 fowlpox vir
35	28	66.7	457	1	ERF1_GIALA	O9ACPI giardia lam
36	28	66.7	496	1	NIFE_RHITO	O9AEP4 rhizobium l
37	28	66.7	496	1	NIFE_RHISN	P55673 rhizobium s
38	28	66.7	530	1	UL21_HSVB	P28972 equine herp
39	28	66.7	624	1	SERA_ARATH	O04130 arabidopsis
40	28	66.7	778	1	DEGL_CAEEL	P24583 caenorhabd1
41	28	66.7	804	1	SCY1_YEAST	P53009 saccharomyc
42	28	66.7	865	1	TOP1_BCOIT	P06612 escherichia
43	28	66.7	868	1	TOP1_PASMU	O9C830 pasteurilla
44	28	66.7	874	1	STA_HAEIN	P43815 haemophilus
45	28	66.7	1502	1	NI70_YEAST	P38181 saccharomyc

ALIGNMENTS

RESULT 1

ID	LDLR_HUMAN	STANDARD	PRT	860 AA.
AC	P01130:			
DT	21-JUL-1986 (rel. 01, Created)			
DT	21-JUL-1986 (rel. 01, Last sequence update)			
DT	16-OCT-2001 (rel. 40, Last annotation update)			
DE	Low-density lipoprotein receptor precursor (LDL receptor).			
GN	LDLR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85024898; PubMed=6091915;			
RA	Yamamoto T., Davis C.G., Brown M.S., Schneider W.J., Casey M.L.,			
RA	Goldstein J.L., Russell D.W.;			
RT	"The human LDL receptor: a cysteine-rich protein with multiple Alu			
RT	sequences in its mRNA.";			
RL	Cell 39:27-38(1984).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85218750; PubMed=2988123;			
RA	Suedhof T.C., Goldstein J.L., Brown M.S., Russell D.W.;			
RT	"The LDL receptor gene: a mosaic of exons shared with different			
RT	proteins.";			
RL	Science 228:815-822(1985).			
RN	[3]			
RP	MUTAGENESIS OF CYTOPLASMIC DOMAIN.			
RX	MEDLINE=87166014; PubMed=3104336;			
RA	Davis C.G., van Driel I.R., Russell D.W., Brown M.S., Goldstein J.L.;			
RT	"The low density lipoprotein receptor. Identification of amino acids			
RT	in cytoplasmic domain required for rapid endocytosis.";			
RL	J. Biol. Chem. 262:4075-4082(1987).			
RN	[4]			
RP	O-GLYCOSYLATION DOMAIN.			
RX	MEDLINE=86140036; PubMed=3005267;			
RA	Davis C.G., Elhammer A., Russell D.W., Schneider W.J., Kornfeld S.,			
RA	Brown M.S., Goldstein J.L.;			
RT	"Deletion of clustered O-linked carbohydrates does not impair			
RT	function of low density lipoprotein receptor in transfected			
RT	fibroblasts.";			
RL	J. Biol. Chem. 261:2828-2838(1986).			
RN	[5]			
RP	STRUCTURE BY NMR OF 20-67.			
RX	MEDLINE=95327641; PubMed=7603991;			
RA	Daly N.L., Scanlon M.J., Djordjevic J.T., Kroon P.A., Smith R.;			
RT	"Three-dimensional structure of a cysteine-rich repeat from the low-			
RT	density lipoprotein receptor.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 92:6334-6338(1995).			
RN	[6]			
RP	STRUCTURE BY NMR OF 65-104.			
RX	MEDLINE=96062511; PubMed=7578052;			
RA	Daly N.L., Djordjevic J.T., Kroon P.A., Smith R.;			
RT	"Three-dimensional structure of the second cysteine-rich repeat from			

RT the human low-density lipoprotein receptor.";
 RL Biochemistry 34:14474-14481(1995).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 196-232.
 RX MEDLINE=97404319; PubMed=9262405;
 RA Fass D., Blacklow S.C., Kim P.S., Berger J.M.;
 RT "Molecular basis of familial hypercholesterolemia from structure of
 RL LDL receptor module.";
 RL Nature 388:691-693(1997).
 RN [8]
 RP REVIEW ON FH VARIANTS.
 RX MEDLINE=93250847; PubMed=1301956;
 RA Hobbs H.H., Brown M.S., Goldstein J.L.;
 RT "Molecular genetics of the LDL receptor gene in familial
 RL hypercholesterolemia.";
 RL Hum. Mutat. 1:445-466(1992).
 RN [9]
 RP REVIEW ON FH VARIANTS.
 RX MEDLINE=97169388; PubMed=9016531;
 RA Varret M., Rabes J.-P., Colod-Beroud G., Junien J., Boileau C.,
 RT "Software and database for the analysis of mutations in the human LDL
 RL receptor gene.";
 RL Nucleic Acids Res. 25:172-180(1997).
 RN [10]
 RP VARIANTS FH AFRIKANER-1; -2 AND -3.
 RX MEDLINE=89340934; PubMed=2569482;
 RA Lettersdorf E., van der Westhuyzen D.R., Coetzee G.A., Hobbs H.H.;
 RT "Two common low density lipoprotein receptor gene mutations cause
 RL familial hypercholesterolemia in Afrikaners.";
 RL J. Clin. Invest. 84:954-961(1989).
 RN [11]
 RP VARIANT FH BARI.
 RX MEDLINE=86161891; PubMed=3955657;
 RA Davys C.G., Lehman M.A., Russell D.W., Anderson R.G.W., Brown M.S.,
 RL Goldstein J.L.;
 RT "The J.D. mutation in familial hypercholesterolemia: amino acid
 RL substitution in cytoplasmic domain impedes internalization of LDL
 RL receptors.";
 RL Cell 45:15-24(1986).
 RN [12]
 RP VARIANTS FH DURBAN-1 AND -2.
 RX MEDLINE=93350003; PubMed=8347689;
 RA Rubinstein D.C., Jialal I., Lettersdorf E., Coetzee G.A.,
 RT van der Westhuyzen D.R.;
 RT "Identification of two new LDL-receptor mutations causing homozygous
 RL familial hypercholesterolemia in a South African of Indian origin.";
 RL Biochim. Biophys. Acta 1182:75-82(1993).
 RN [13]
 RP VARIANTS FH FRENCH CANADIAN-2; -3 AND -4.
 RX MEDLINE=90203205; PubMed=2318961;
 RA Lettersdorf E., Tobin E.J., Davignon J., Hobbs H.H.;
 RT "Common low-density lipoprotein receptor mutations in the French
 RL Canadian population.";
 RL J. Clin. Invest. 85:1014-1023(1990).
 RN [14]
 RP VARIANT FH OSAKA-3.
 RX MEDLINE=93076772; PubMed=1446662;
 RA Miyake Y., Tajima S., Funahashi T., Yamamura T., Yamamoto A.;
 RT "A point mutation of low-density lipoprotein receptor causing rapid
 RL degradation of the receptor.";
 RL Eur. J. Biochem. 210:1-7(1992).
 RN [15]
 RP VARIANT FH PISCATAMAY.
 RX MEDLINE=91328152; PubMed=1867200;
 RA Weiner V., Landsberger D., Berkman N., Reshef A., Segal P.,
 RA Seftel H.C., van der Westhuyzen D.R., Jeenanah M.S., Coetzee G.A.,
 RL Lettersdorf E.;
 RT "A common lithuanian mutation causing familial hypercholesterolemia
 RL in Ashkenazi Jews.";
 RL Am. J. Hum. Genet. 49:443-449(1991).
 RN [16]
 RP VARIANT FH SAFED.

RX MEDLINE=93216279; PubMed=8462973;
 RA Lettersdorf E., Reshef A., Weiner V., Dann E.J., Beigel Y.,
 RA van Rogen F.G., van der Westhuyzen D.R., Coetzee G.A.;
 RT "A missense mutation in the low density lipoprotein receptor gene
 RL causes familial hypercholesterolemia in Sephardic Jews.";
 RL Hum. Genet. 91:141-147(1993).
 RN [17]
 RP VARIANT FH TRIESTE.
 RX MEDLINE=94224219; PubMed=8166830;
 RA Lelli N., Garuti R., Pedrazzi P.,
 RA Tiozzo R., Catlin L., Valenti M., Rolletti M., Bertolini S.,
 RA Stefanutti C., Calandra S.;
 RT "A new missense mutation (Cys297-->Phe) of the low density
 RL lipoprotein receptor in Italian patients with familial
 RL hypercholesterolemia (FHTrieste).";
 RL Hum. Genet. 93:538-540(1994).
 RN [18]
 RP VARIANT FH ZAMBIA.
 RX MEDLINE=89264579; PubMed=2726768;
 RA Soutar A.K., Knight B.L., Patel D.D.;
 RT "Identification of a point mutation in growth factor repeat C of the
 RL low density lipoprotein-receptor gene in a patient with homozygous
 RL familial hypercholesterolemia that affects ligand binding and
 RL intracellular movement of receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4166-4170(1989).
 RN [19]
 RP VARIANT FH ZAMBIA.
 RX MEDLINE=93100522; PubMed=1464748;
 RA Rubinstein D.C., Coetzee G.A., Marais A.D., Lettersdorf E.,
 RA Seftel H.C., van der Westhuyzen D.R.;
 RT "Identification and properties of the proline664-leucine mutant LDL
 RL receptor in South Africans of Indian origin.";
 RL J. Lipid Res. 33:1647-1653(1992).
 RN [20]
 RP VARIANTS FH PORI HIS-401 AND TURKU ASP-844.
 RX MEDLINE=96029270; PubMed=7573037;
 RA Koivisto V.-M., Viikari J.S., Kontula K.;
 RT "Molecular characterization of minor gene rearrangements in Finnish
 RL patients with heterozygous familial hypercholesterolemia:
 RL identification of two common missense mutations (Gly823-->Asp and
 RL Leu380-->His) and eight rare mutations of the LDL receptor gene.";
 RL Am. J. Hum. Genet. 57:789-797(1995).
 RN [21]
 RP VARIANTS FH LYS-140; SER-338 AND LEU-685.
 RX MEDLINE=96011600; PubMed=7583548;
 RA Maruyama T., Miyake Y., Tajima S., Harada-Shiba M., Yamamura T.,
 RA Tsuchino M., Kishino B.-I., Horiguchi Y., Funahashi T., Matsuzawa Y.,
 RA Yamamoto A.;
 RT "Common mutations in the low-density lipoprotein-receptor gene
 RL causing familial hypercholesterolemia in the Japanese population.";
 RL Arterioscler. Thromb. Vasc. Biol. 15:1713-1718(1995).
 RN [22]
 RP VARIANT FH FRENCH HIS-564.
 RX MEDLINE=96055524; PubMed=7550239;
 RA Triot-Guerber F., Saint-Jore B., Valenti K., Foulon T., Bost M.,
 RA Hadjian A.J.;
 RT "Identification of a mutation, N543H, in exon 11 of the low-density
 RL lipoprotein receptor gene in a French family with familial
 RL hypercholesterolemia.";
 RL Hum. Mutat. 6:87-88(1995).
 RN [23]
 RP VARIANTS FH SWEDISH LYS-277; THR-423 AND ASN-579.
 RX MEDLINE=95362239; PubMed=7635461;
 RA Ekstrom U., Adrhamson M., Sveger T., Lombardi P., Nilsson-Ehle P.;
 RT "An efficient screening procedure detecting six novel mutations in
 RL the LDL receptor gene in Swedish children with
 RL hypercholesterolemia.";
 RL Hum. Genet. 96:147-150(1995).
 RN [24]
 RP VARIANT FH NORWEGIAN ASN-487 DEL.
 RX MEDLINE=95362260; PubMed=7635482;
 RA Ieren T.P., Solberg K., Rodnighi O.K., Tonstad S., Ose L.;
 RT "Two novel point mutations in the EGF precursor homology domain of

Query Match 78.6%; Score 33; DB 1; Length 860;
 Best Local Similarity 87.5%; Pred. No. 19;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SEVNLLAE 8
 DB 637 SDVNLAE 644

RESULT 2

NIFE_RHIME
 ID NIFE_RHIME STANDARD: PRT: 476 AA.
 AC 092ZT2;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DE NITROGENASE iron-polyhydnum cofactor biosynthesis protein nife.
 GN NIFE OR RA0452 OR SMA0830.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSyma (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=1021;
 RX MEDLINE=21396509; PubMed=11481432;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barilay-Hubler F., Bowser L., Capela D., Galbert F., Gouy J.,
 RA Gurjal M., Hong A., Hutzar L., Hyman R.W., Kahn M.L.,
 RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti pSyma megaplasmid."
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 CC -1- FUNCTION: THIS PROTEIN MAY PLAY A ROLE IN THE BIOSYNTHESIS OF
 CC THE PROSTHETIC GROUP OF NITROGENASE (FEMO COFACTOR).
 CC -1- PATHWAY: FE-MO COFACTOR BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE NIFD/NIFK/NIFE/NIFN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AE007236; AAK65110.1;
 DR PROSITE: PS00699; NITROGENASE_1.1; 1.
 DR PROSITE: PS00090; NITROGENASE_1.2; 1.
 KW Nitrogen fixation; Plasmid; Complete proteome.
 SQ SEQUENCE 476 AA; 52358 MW; 83BCE94B4B14C713 CRC64;

Query Match 76.2%; Score 32; DB 1; Length 476;
 Best Local Similarity 55.6%; Pred. No. 17;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAE 9
 DB 195 SDVNLAE 203

RESULT 3

HEX_ADEG1 STANDARD: PRT: 942 AA.
 AC P42671; 064758;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hexon protein (late protein 2).
 DB P11.

OS Avian adenovirus gall (strain Phe1ps) (Fowl adenovirus 1) (CELO).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
 OX NCBI_TaxID=10553;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=96186720; PubMed=8627769;
 RA Chlocca S., Kurzbauer R., Schaffner G., Baker A., Mautner V.,
 RA Cotten M.;
 RT "The complete DNA sequence and genomic organization of the avian
 RT adenovirus CELO."
 RL J. Virol. 70:2939-2949(1996).
 RN [2]
 RP SEQUENCE OF 913-942 FROM N.A.

RX MEDLINE=93362429; PubMed=8395124;
 RA Cai F., Weber J.M.;
 RT "Organization of the avian adenovirus genome and the structure of its
 RT endonuclease."
 RT Virology 196:358-362(1993).
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE STRUCTURAL PROTEINS IN THE
 CC VIRAL COAT AND IS SYNTHESIZED DURING LATE INFECTION (BY
 CC SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: U46933; AAC54912.1;
 DR EMBL: L13161; AAA51401.1;
 DR HSSP: P03327; IDHX.
 DR InterPro: IPR00736; Adeno_hexon.
 DR Pfam: PF01065; Adeno_hexon; 1.
 DR ProDom: PD002815; Adeno_hexon; 1.
 KW Coat protein; Hexon protein; Late protein.
 SQ SEQUENCE 942 AA; 106709 MW; 7F4CE8D5F17D051B CRC64;

Query Match 76.2%; Score 32; DB 1; Length 942;
 Best Local Similarity 66.7%; Pred. No. 36;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAE 9
 DB 581 TEVNLMANF 589

RESULT 4

ORN_VIBCH STANDARD: PRT: 181 AA.
 ID ORN_VIBCH
 AC 09KV17;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oligoribonuclease (EC 3.1.-.-).
 GN ORN OR VC0341.

OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibri.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio

```

RT cholerae."
RL Nature 406:477-483(2000).
CC -1- FUNCTION: 3'-TO-5' EXORIBONUCLEASE SPECIFIC FOR SMALL
CC OLIGORIBONUCLEOTIDES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE OLIGORIBONUCLEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004123; AAF93514.1; -.
DR TIGR: VC0341; -.
DR InterPro: IPR000520; Exonuclease.
DR Pfam: PF00929; Exonuclease; 1.
DR SMART: SM00479; EXOIII; 1.
DR Hydrolase: Exonuclease; Complete proteome.
FT ACT_SITE 129 129 POTENTIAL.
SQ SEQUENCE 181 AA; 20929 MW; AAC071B9C94FEFE3 CRC64;

Query Match 73.8%; Score 31; DB 1; Length 181;
Best Local Similarity 75.0%; Fred. No. 9.9;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SEVNLAE 8
DB 35 SEVNLAE 42

RESULT 5
NEK2_MOUSE STANDARD; PRT; 443 AA.
AC 035942: 035959;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Serine/threonine-protein kinase NEK2 (EC 2.7.1.1.-) (NIMA-related
DE protein kinase 2).
GN NEK2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
[1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC STRAIN=SWISS WEBSTER; TISSUE=Testis;
RX MEDLINE=97330684; PubMed=9187143;
RA Rhee K., Wolgemuth D.J.;
RT "The NIMA-related kinase 2, Nek2, is expressed in specific stages of
RT the meiotic cell cycle and associates with meiotic chromosomes.";
RL Development 124:2167-2177(1997).
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=98243037; PubMed=9583679;
RA Arama E., Janai A., Killin G., Motro B.;
RT "Murine NIMA-related kinases are expressed in patterns suggesting
RT distinct functions in gametogenesis and a role in the nervous
RT system.";
RL Oncogene 16:1813-1823(1998).
RN
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Brain;
RX MEDLINE=98096235; PubMed=9434622;
RA Tanaka K., Parvinen M., Nigg E.A.;
RT "The in vivo expression pattern of mouse Nek2, a NIMA-related kinase,
RT indicates a role in both mitosis and meiosis.";
RL Exp. Cell Res. 237:264-274(1997).
CC -1- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION.
CC -1- MAY HAVE A ROLE AT THE G2-M TRANSITION. MAY ALSO PLAY A ROLE IN

```

```

CC MEIOSIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: MOST ABUNDANTLY EXPRESSED IN TESTIS. LOW
CC LEVELS FOUND IN MID-GESTATION EMBRYO, OVARY, PLACENTA, INTESTINE,
CC THYMUS AND SKIN. WITHIN THE TESTIS, EXPRESSION RESTRICTED TO GERM
CC CELLS WITH HIGHEST LEVELS DETECTED IN SPERMATOCYTES AT PACHYENE
CC AND DIPTERENE STAGES. ALSO EXPRESSED IN MEIOTIC PACHYENE OOCYTES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIMA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U95610; AAB67973.1; -.
DR EMBL: AF013166; AAC35393.1; -.
DR EMBL: AF007247; AAB70470.1; -.
DR HSSP: P24941; ICKP.
DR MGD: MGI:109359; NEK2.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP_FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR TRANSFERASE: Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Nuclear protein; Phosphorylation; Cell cycle; Cell division; Meiosis.
FT DOMAIN 8 271 PROTEIN KINASE.
FT NR_BIND 14 22 ATP (BY SIMILARITY).
FT BINDING 37 37 ATP (BY SIMILARITY)..
FT ACT_SITE 141 141 BY SIMILARITY.
FT MOD_RES 179 179 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 443 AA; 51307 MW; DE09565C378307E1 CRC64;

Query Match 73.8%; Score 31; DB 1; Length 443;
Best Local Similarity 87.5%; Fred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAE 8
DB 54 SEVNLAE 61

RESULT 6
NEK2_HUMAN STANDARD; PRT; 445 AA.
AC P51955;
DT 01-OCT-1996 (Rel. 34; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Serine/threonine-protein kinase NEK2 (EC 2.7.1.1.-) (NIMA-related
DE protein kinase 2) (NIMA-like protein kinase 1) (HSK 21).
GN NEK2 OR NTK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell, Placenta, and Nasopharynx;
RX MEDLINE=94368699; PubMed=7522034;
RA Schultz S.J., Fry A.M., Suetterlin C., Ried T., Nigg E.A.;
RT "Cell cycle-dependent expression of Nek2, a novel human protein
RT nidulans.";
RL Cell Growth Differ. 5:625-635(1994).
RN
RP SEQUENCE FROM N.A.

```

```

RA Lu K.P., Hunter T.;
RT "Molecular cloning and expression of NUK1, a human NIMA-like kinase.";
RN Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE OF 83-203 FROM N.A.
RX MEDLINE-94100173; PubMed-8274451;
RA Schultz S.J., Migg E.A.;
RT "Identification of 21 novel human protein kinases, including 3 members
of a family related to the cell cycle regulator nima of Aspergillus
nidulans.";
RL Nature 392:353-358(1998).
CC -1- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION.
CC MAY HAVE A ROLE AT THE G2-M TRANSITION. MAY ALSO PLAY A ROLE IN
CC MEIOSIS.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES THROUGHOUT S PHASE AND SHOWS
CC MAXIMAL LEVELS IN LATE G2. THIS EXPRESSION PATTERN IS HIGHLY
CC REMINISCENT OF THAT OF A AND B CYCLINS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIMA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z29066; CA82309.1; -;
DR EMBL; U11050; AAA19558.1; -;
DR EMBL; Z25425; CA80912.1; -;
DR HSSP; P00518; 1PHK.
DR MIM; 604043; -;
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_Thr_Pkinase.
DR Pfam; PF00069; pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Nuclear protein; Phosphorylation; Cell cycle; Cell division; Meiosis.
FT DOMAIN 8 271 PROTEIN KINASE.
FT NP_BIND 14 22 ATP (BY SIMILARITY).
FT BINDING 37 37 ATP (BY SIMILARITY).
FT ACT_SITE 141 141 BY SIMILARITY.
FT MOD_RES 179 179 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 84 85 IV -> LY (IN REF. 3).
SQ SEQUENCE 445 AA; 51763 MW; D33A3778ABBD9E CRC64;

Query Match 73.8%; Score 31; DB 1; Length 445;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAE 8
Db 54 SEVNLAE 61
|||||
|

RESULT 7
Y041_AQUAE STANDARD; PRT; 103 AA.
ID Y041_AQUAE
AC 066457;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AQ_041.
GN AQ_041.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID-63363;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE-98196666; PubMed-9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aulay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
CC -1- SIMILARITY: BELONGS TO THE UPF0102 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000671; AAC06429.1; -;
DR InterPro: IPR003509; UPF0102.
DR Pfam; PF02021; UPF0102; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 103 AA; 11870 MW; 78E285829D39FB7E CRC64;

Query Match 71.4%; Score 30; DB 1; Length 103;
Best Local Similarity 62.5%; Pred. No. 9;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 EVNLAEF 9
Db 32 EIDLAEF 39
|||||
|

RESULT 8
Y778_AQUAE STANDARD; PRT; 328 AA.
ID Y778_AQUAE
AC 066974;
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Hypothetical protein AQ_778.
GN AQ_778.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID-63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE-98196666; PubMed-9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aulay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
CC -1- SIMILARITY: BELONGS TO THE UPF0052 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000707; AAC06942.1; -;
DR InterPro: IPR000873; AMP_bind.
DR InterPro: IPR002882; UPF0052.
DR Pfam; PF01933; UPF0052; 1.
DR PROSITE; PS00455; AMP_BINDING; UNKNOWN_1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 328 AA; 35957 MW; 30BB21AE99558F6B CRC64;

```

```

Query Match          71.4%; Score 30; DB 1; Length 328;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      3 VNLAEEF 9
      |||:|
DB      133 VNLAEEF 139

RESULT 9
ID_YD15_SCHPO      STANDARD; PRT; 581 AA.
AC  Q10238;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  01-OCT-1996 (Rel. 34, Last annotation update)
DE  Putative regulatory protein CAG9.05.
GN  SPAC49.05.
OS  Schizosaccharomyces pombe (fission yeast).
OC  Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC  Schizosaccharomycetales; Schizosaccharomycetaceae;
OC  Schizosaccharomyces.
OX  NCBI_TaxID=4896;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=972;
RA  Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL  Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: SOME, TO DROSOPHILA PUM.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; 269727; CAA93555.1; -.
DR  InterPro; IPR001313; PDM.
DR  Pfam; PF00806; PUF; 6.
DR  SMART; SM00025; Pumilio; 7.
KW  Hypothetical protein.
SQ  SEQUENCE 581 AA; 66666 MW; 93A360BA7BD0A40E CRC64;

Query Match          71.4%; Score 30; DB 1; Length 581;
Best Local Similarity 62.5%; Pred. No. 59;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY      1 SEVNLAEE 8
      ||:|:|:|
DB      547 SQINLAEE 554

RESULT 10
MEC4_CAEEL      STANDARD; PRT; 768 AA.
ID  MEC4_CAEEL
AC  P24612;
DT  01-MAR-1992 (Rel. 21, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Mechanosensory protein 4 (Degenerin mec-4).
GN  MEC-4 OR MEC-13 OR TOLC8.7.
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC  Rhabditidae; Peloderinae; Caenorhabditis.
OX  NCBI_TaxID=6239;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-BRISTOL N2;
RX  MEDLINE=96251674; PubMed=8655580;
RA  Lai C.C., Hong K., Kinnell M., Chalfie M.;

```

```

RT  "Sequence and transmembrane topology of MEC-4, an ion channel subunit
RT  required for mechanotransduction in Caenorhabditis elegans.";
RL  J Cell Biol. 133:1071-1081(1996).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-BRISTOL N2;
RA  Wohlmann P., Hawkins J.;
RL  Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE OF 271-768 FROM N.A.
RC  STRAIN-BRISTOL N2;
RX  MEDLINE=91156026; PubMed=1672038;
RA  Driscoll M., Chalfie M.;
RT  "The mec-4 gene is a member of a family of Caenorhabditis elegans
RT  genes that can mutate to induce neuronal degeneration.";
RL  Nature 349:588-593(1991).
CC  -1- FUNCTION: PROBABLE SODIUM CHANNEL SUBUNIT. MAY BE NEEDED FOR
CC  MECHANOSENSORY TRANSDUCTION (TOUCH SENSITIVITY).
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC  -1- DISEASE: MUTATIONS IN MEC-4 RESULTS IN THE DEGENERATION OF A SMALL
CC  SET OF NEURONS WHICH TYPICALLY SMELL TO SEVERAL TIME THEIR NORMAL
CC  DIAMETER BEFORE THEY DISAPPEAR. PRESUMABLY DUE TO LYSIS.
CC  -1- SIMILARITY: BELONGS TO THE AMILORIDE-SENSITIVE SODIUM CHANNELS
CC  FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U53669; AAC47265.1; -.
DR  EMBL; U58726; AAB00580.1; ALT_INIT.
DR  EMBL; X58982; CAA41731.1; -.
DR  PIR; S13645; S13645.
DR  WormPep; T01C8.7; CE07463.
DR  InterPro; IPR001873; ASC.
DR  Pfam; PF00858; ASC; 1.
DR  PRINTS; PR01078; AMINACHANNEL.
DR  PROSITE; PS01206; ASC; 1.
KW  Ionic channel; Transmembrane; Ion transport; Glycoprotein;
KW  Neurodegeneration.
SQ  SEQUENCE 768 AA; 87194 MW; 8899D0FCB7F82C6 CRC64;

Query Match          71.4%; Score 30; DB 1; Length 768;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      3 VNLAEEF 9
      |||:|
DB      709 VNLAEEF 715

RESULT 11
MEC4_CAEER      STANDARD; PRT; 769 AA.
ID  MEC4_CAEER

```

```

AC Q17296;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Mechanosensory protein 4 (Degenerin mec-4).
GN MEC-4.
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96251674; PubMed=86555580;
RA Lai C.C., Hong K., Kinnell M., Chalfie M., Driscoll M.;
RT "Sequence and transmembrane topology of MEC-4, an ion channel subunit
RT required for mechanotransduction in Caenorhabditis elegans.";
RL J. Cell Biol. 133:1071-1081(1996).
CC -1- FUNCTION: PROBABLE SODIUM CHANNEL SUBUNIT. MAY BE NEEDED FOR
CC MECHANOSENSORY TRANSDUCTION (TOUCH SENSITIVITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE AMILORIDE-SENSITIVE SODIUM CHANNELS
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U53670; AAC47264.1; -.
DR InterPro: IPR001873; ASC.
DR Pfam: PF00858; ASC. 1.
DR PRINTS: PR01078; AMINACHANNEL.
DR PROSITE: PS01206; ASC. 1.
KW Ionic channel; Transmembrane; Ion transport; Glycoprotein;
KW Neurodegeneration.
FT DOMAIN 1 109 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 110 130 POTENTIAL.
FT DOMAIN 131 719 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 720 740 POTENTIAL.
FT DOMAIN 741 769 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 485 485 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 769 AA; 87141 MW; 430BB718BA0F5002 CRC64;
Query Match 71.4%; Score 30; DB 1; Length 769;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 3 VNLLAEF 9
DB 710 VNLLADF 716

```

```

OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92407468; PubMed=1527478;
RA Bishop R.W.;
RT "Structure of the hamster low density lipoprotein receptor gene.";
RL J. Lipid Res. 33:549-557(1992).
RN [2]
RP SEQUENCE OF 570-615 FROM N.A.
RX MEDLINE=87064645; PubMed=3785227;
RA Sege R.D., Kozarsky K.F., Krieger M.;
RT "Characterization of a family of gamma-ray-induced CHO mutants
RT demonstrates that the LDL locus is diploid and encodes the
RT low-density lipoprotein receptor.";
RL Mol. Cell. Biol. 6:3268-3277(1986).
CC -1- FUNCTION: BINDS LDL, THE MAJOR CHOLESTEROL-CARRYING LIPOPROTEIN
CC OF PLASMA, AND TRANSPORTS IT INTO CELLS BY ENOCYTOSIS. IN ORDER
CC TO BE INTERNALIZED, THE RECEPTOR-LIGAND COMPLEXES MUST FIRST
CC CLUSTER INTO CLATHRIN-COATED PITS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: CONTAINS 7 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M94387; AA51449.1; -.
DR EMBL: M18677; -. NOT_ANNOTATED_CDS.
DR PIR: A48908; ORHYLD.
DR HSSP: P01130; ILDR.
DR InterPro: IPR000152; Asx hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-like.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR000033; LDL_receptor_rep.
DR Pfam: PF00008; EGF. 3.
DR Pfam: PF00057; Idl_recept_a; 7.
DR Pfam: PF00058; Idl_recept_b; 5.
DR PRINTS: PR00261; LDLRECEPTOR.
DR SMART: SM00179; EGF_CA; 1.
DR SMART: SM00192; EGF_Like; 2.
DR SMART: SM00192; LDLA; 7.
DR PROSITE: PS000135; LV; 5.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS01209; LDLRA_1; 7.
DR PROSITE: PS50068; LDLRA_2; 7.
KW Glycoprotein; LDL; Cholesterol metabolism; Lipid transport;
KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;
KW EGF-like domain; Repeat.
FT SIGNAL 1 21
FT CHAIN 22 854
FT DOMAIN 22 782
FT TRANSMEM 783 804
FT DOMAIN 805 854
FT DOMAIN 25 316
FT FT 25
FT FT 65
FT DOMAIN 66 106
FT DOMAIN 107 145
FT DOMAIN 146 186
FT DOMAIN 196 234
FT DOMAIN 235 273
FT DOMAIN 275 314
FT DOMAIN 315 354

```

```

FT DOMAIN 355 394 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 398 439 LDL-RECEPTOR CLASS B 1.
FT REPEAT 440 485 LDL-RECEPTOR CLASS B 2.
FT REPEAT 486 528 LDL-RECEPTOR CLASS B 3.
FT REPEAT 529 572 LDL-RECEPTOR CLASS B 4.
FT REPEAT 573 615 LDL-RECEPTOR CLASS B 5.
FT REPEAT 616 657 LDL-RECEPTOR CLASS B 6.
FT REPEAT 663 712 EGF-LIKE 3.
FT DOMAIN 721 765 CLUSTERED O-LINKED OLIGOSACCHARIDES.
FT SITE 817 822 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 27 39 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 68 82 BY SIMILARITY.
FT DISULFID 75 95 BY SIMILARITY.
FT DISULFID 89 104 BY SIMILARITY.
FT DISULFID 109 121 BY SIMILARITY.
FT DISULFID 116 134 BY SIMILARITY.
FT DISULFID 128 143 BY SIMILARITY.
FT DISULFID 148 160 BY SIMILARITY.
FT DISULFID 155 173 BY SIMILARITY.
FT DISULFID 167 184 BY SIMILARITY.
FT DISULFID 198 210 BY SIMILARITY.
FT DISULFID 205 223 BY SIMILARITY.
FT DISULFID 217 232 BY SIMILARITY.
FT DISULFID 237 249 BY SIMILARITY.
FT DISULFID 244 262 BY SIMILARITY.
FT DISULFID 256 271 BY SIMILARITY.
FT DISULFID 277 290 BY SIMILARITY.
FT DISULFID 285 303 BY SIMILARITY.
FT DISULFID 297 314 BY SIMILARITY.
FT DISULFID 319 330 BY SIMILARITY.
FT DISULFID 326 339 BY SIMILARITY.
FT DISULFID 341 353 BY SIMILARITY.
FT DISULFID 359 369 BY SIMILARITY.
FT DISULFID 365 378 BY SIMILARITY.
FT DISULFID 380 393 BY SIMILARITY.
FT DISULFID 667 681 BY SIMILARITY.
FT DISULFID 677 696 BY SIMILARITY.
FT DISULFID 698 711 BY SIMILARITY.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 657 657 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 854 AA; 94516 MW; 5D50EBFBD834B35E CRC64;

Query Match 71.4%; Score 30; DB 1; Length 854;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SEVNLAE 8
DB 637 SDVNLVAE 644

RESULT 13
LDLR_MOUSE STANDARD; PRT; 864 AA.
ID LDLR_MOUSE
AC P35951;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Low-density lipoprotein receptor precursor (LDL receptor).
GN LDLR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=Liver;
RX MEDLINE=93221531; PubMed=8466528;
RA Hofter M.J.V., van Eck M.M., Petit J.F., van der Zee A., de Wit E.,
RA Meijer D., Grosveld G., Havekes L.M., Hofker M.H., Frans R.R.;

```

```

RT "The mouse low density lipoprotein receptor gene: cDNA sequence and
RT exon-intron structure."
RT Biochem. Biophys. Res. Commun. 191:880-886(1993).
RL [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93117934; PubMed=1475710;
RA Polyano W.J., Dichek D.A., Mason J., Anderson W.F.;
RT "Molecular cloning and nucleotide sequence of cDNA encoding a
RT functional murine low-density-lipoprotein receptor."
RT Somat. Cell Mol. Genet. 18:443-450(1992).
RL -1- FUNCTION: BINDS LDL, THE MAJOR CHOLESTEROL-CARRYING LIPOPROTEIN
CC OF PLASMA, AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS. IN ORDER
CC TO BE INTERNALIZED, THE RECEPTOR-LIGAND COMPLEXES MUST FIRST
CC CLUSTER INTO CLATHRIN-COATED PITS.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: CONTAINS 7 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z19521; CAAT9581.1; -
DR EMBL: X64414; CAAT4759.1; -
DR PIR: JN0461; ORMSLD.
DR HSSP: P01130; ILDR.
DR MGD: MGI:96765; Ldlr.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-Like.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR000033; Ldl_recept_rep.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00057; ldl_recept_a; 7.
DR Pfam: PF00058; ldl_recept_b; 5.
DR PRINTS: PR00261; LDLRECEPTOR.
DR SMART: SM00179; EGF_CA; 1.
DR SMART: SM00001; EGF_Like; 2.
DR SMART: SM00192; LDLA; 7.
DR SMART: SM00135; LY; 5.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS01209; LDLRA_1; 7.
DR PROSITE: PS50068; LDLRA_2; 7.
DR PROSITE: LDL; Cholesterol metabolism; Lipid transport;
KW Glycoprotein; LDL; Cholesterol metabolism; Receptor; Signal;
KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;
KW EGF-like domain; Repeat.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 864 LOW-DENSITY LIPOPROTEIN RECEPTOR.
FT DOMAIN 22 790 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 791 812 POTENTIAL.
FT DOMAIN 813 864 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 65 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 66 106 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 107 145 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 146 186 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 196 234 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 235 273 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 275 314 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 315 354 EGF-LIKE 1.
FT DOMAIN 355 394 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 398 439 LDL-RECEPTOR CLASS B 1.
FT REPEAT 440 485 LDL-RECEPTOR CLASS B 2.
FT REPEAT 486 528 LDL-RECEPTOR CLASS B 3.
FT REPEAT 529 572 LDL-RECEPTOR CLASS B 4.

```



```

FT REPEAT 573 615 LDL-RECEPTOR CLASS B 5.
FT REPEAT 616 657 LDL-RECEPTOR CLASS B 6.
FT DOMAIN 663 713 EGF-LIKE 3.
FT SITE 722 770 CLUSTERED O-LINKED OLIGOSACCHARIDES.
FT SITE 827 832 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 68 82 BY SIMILARITY.
FT DISULFID 75 95 BY SIMILARITY.
FT DISULFID 89 104 BY SIMILARITY.
FT DISULFID 109 121 BY SIMILARITY.
FT DISULFID 116 134 BY SIMILARITY.
FT DISULFID 128 143 BY SIMILARITY.
FT DISULFID 148 160 BY SIMILARITY.
FT DISULFID 155 173 BY SIMILARITY.
FT DISULFID 167 184 BY SIMILARITY.
FT DISULFID 198 210 BY SIMILARITY.
FT DISULFID 205 223 BY SIMILARITY.
FT DISULFID 217 232 BY SIMILARITY.
FT DISULFID 237 249 BY SIMILARITY.
FT DISULFID 244 262 BY SIMILARITY.
FT DISULFID 256 271 BY SIMILARITY.
FT DISULFID 277 290 BY SIMILARITY.
FT DISULFID 285 303 BY SIMILARITY.
FT DISULFID 297 314 BY SIMILARITY.
FT DISULFID 319 330 BY SIMILARITY.
FT DISULFID 326 339 BY SIMILARITY.
FT DISULFID 341 353 BY SIMILARITY.
FT DISULFID 359 369 BY SIMILARITY.
FT DISULFID 365 378 BY SIMILARITY.
FT DISULFID 380 393 BY SIMILARITY.
FT DISULFID 667 682 BY SIMILARITY.
FT DISULFID 678 697 BY SIMILARITY.
FT DISULFID 699 712 BY SIMILARITY.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 23 23 V -> A (IN REF. 2).
FT CONFLICT 27 27 G -> C (IN REF. 2).
FT CONFLICT 61 61 K -> Q (IN REF. 2).
FT CONFLICT 144 144 P -> Q (IN REF. 2).
FT CONFLICT 156 156 K -> N (IN REF. 2).
FT CONFLICT 178 178 H -> D (IN REF. 2).
FT CONFLICT 186 187 AE -> GR (IN REF. 2).
FT CONFLICT 820 821 MISSING (IN REF. 2).
SQ SEQUENCE 864 AA: 95134 MW: 4AE15953ED11E0E CRC64:

Query Match
Best Local Similarity 71.4%: Score 30; DB 1; Length 864;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SEVNLAE 8
Db 637 SDVNLVAE 644

RESULT 14
ID IF2_CHLPN STANDARD: PRT: 890 AA.
AC Q928M1: Q9JRX1:
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Translation initiation factor IF-2.
GN INF8 OR CPN0317 OR CP0440.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
Kalmán S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,

```

```

RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RN Nat. Genet. 21:385-389(1999).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson M., Deroy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39.";
RN Nucleic Acids Res. 28:1397-1406(2000).
RP [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shital M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA";
RN Nucleic Acids Res. 28:2311-2314(2000).
RP [1]
RP FUNCTION: IF-2, ONE OF THE ESSENTIAL COMPONENTS FOR THE INITIATION
OF PROTEIN SYNTHESIS IN VITRO, PROTECTS FORMYL METHIONYL-TRNA FROM
SPONTANEOUS HYDROLYSIS AND PROMOTES ITS BINDING TO THE 30S
RIBOSOMAL SUBUNIT. IT IS ALSO INVOLVED IN THE HYDROLYSIS OF GTP
DURING THE FORMATION OF THE 70S RIBOSOMAL COMPLEX (BY SIMILARITY).
CC -! SUBCELLULAR LOCATION: Cytoplasmic.
CC -! SIMILARITY: BELONGS TO THE IF-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001616; AADI8466.1; -
DR EMBL: AE002205; AAF38279.1; -
DR EMBL: AP002546; BAA98527.1; -
DR HSSP: P02990; IFTU.
DR TIGR: CP0440; -
DR InterPro: IPR000795; GTP_EFTU.
DR InterPro: IPR004161; GTP_EFTU_D2.
DR InterPro: IPR000178; IF2.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 2.
DR PRINTS: PR00315; ELONGATNFCF.
DR PRODOM: PD186100; IF2; 1.
DR PROSITE: PS01176; IF2; 1.
KW Initiation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT DOMAIN 398 546 G-DOMAIN.
FT NP_BIND 404 411 GTP (BY SIMILARITY).
FT NP_BIND 450 454 GTP (BY SIMILARITY).
FT NP_BIND 504 507 GTP (BY SIMILARITY).
FT CONFLICT 321 321 T -> A (IN REF. 1).
SQ SEQUENCE 890 AA: 97104 MW: BB84C2C269268781 CRC64:

Query Match
Best Local Similarity 71.4%: Score 30; DB 1; Length 890;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAE 8
Db 521 SEINLPE 528

RESULT 15
Y549_RICPR

```

```

ID Y549_RICPR STANDARD; PRT; 135 AA.
AC 092D01;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein RP549.
GN RP549.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_Taxid=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,
RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RL Nature 396:133-140(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ235272; CAIL498.1;
CC DR Hypothetical protein; Complete proteome.
CC KW SEQUENCE 135 AA; 16051 MW; 21C8CB79EF2127C6 CRC64;
SQ
Query Match 69.0%; Score 29; DB 1; Length 135;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 EVNIIAE 8
|:|:|:|:|
Db 33 ELNIIAE 39

```

Search completed: October 30, 2002, 12:27:54
 Job time : 2.77396 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:23:30 ; Search time 1.72482 seconds
(without alignments)
501.389 Million cell updates/sec

Title: US-09-724-571-91

Perfect score: 42

Sequence: 1 SEVNLLAEF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	81.0	529	T42584	tegument protein 4
2	33	78.6	668	B96999	methy1-accepting c
3	33	78.6	860	QRHULD	LDL receptor precu
4	32	76.2	132	G72351	conserved hypotnet
5	32	76.2	279	T10882	1-aminocyclopropan
6	32	76.2	422	T09120	basic leucine z1pp
7	32	76.2	467	T10854	1-aminocyclopropan
8	32	76.2	472	T10889	1-aminocyclopropan
9	32	76.2	476	D95318	N1fe oxidoreductas
10	31	73.8	181	E82333	oligoribonuclease
11	31	73.8	225	T17795	hypothetical prote
12	31	73.8	236	AD2830	conserved hypotnet
13	31	73.8	240	D97588	hypothetical prote
14	31	73.8	394	2	hypothetical prote
15	31	73.8	411	T15020	hypothetical prote
16	31	73.8	435	T39719	probable porphyrin
17	31	73.8	445	G01452	beta transducin -
18	31	73.8	461	A27672	hypothetical gag p
19	31	73.8	3169	T00265	toxin B - Escheric
20	30	71.4	103	T70303	hypothetical prote
21	30	71.4	133	E86796	arsenate reductase
22	30	71.4	234	AG1167	hypothetical prote
23	30	71.4	328	C70368	conserved hypotnet
24	30	71.4	334	AC3336	thiosulfate-bindin
25	30	71.4	342	AG1851	hypothetical prote
26	30	71.4	446	B83033	probable MFS trans
27	30	71.4	457	A95194	cell division prot
28	30	71.4	457	F98060	cell division prot
29	30	71.4	475	AF0687	bacteriophage tail

30	30	71.4	483	2	A67583	peptidoglycan bind
31	30	71.4	581	2	T38864	probable regulator
32	30	71.4	694	2	A12492	hypothetical prote
33	30	71.4	749	2	T29859	mechanosensory pro
34	30	71.4	854	1	QRHULD	LDL receptor precu
35	30	71.4	862	1	ORMSLD	LDL receptor precu
36	30	71.4	871	2	AE0270	DNA topoisomerase
37	30	71.4	890	2	E86530	initiation factor-
38	30	71.4	890	2	E81576	translation initia
39	30	71.4	890	2	T02083	H+-transporting AT
40	30	71.4	949	2	H85023	hypothetical prote
41	30	71.4	1216	2	T13930	tripeptidyl-peptid
42	30	71.4	1354	2	D71659	prophage p13 prote
43	29	69.0	135	2	D71659	hypothetical prote
44	29	69.0	135	2	D71659	ribosomal protein
45	29	69.0	230	2	S73229	

ALIGNMENTS

RESULT 1

T42584

tegument protein 40 - equine herpesvirus 4 (strain NS80567)

C:Species: equine herpesvirus 4

A:Variety: strain NS80567

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000

C:Accession: T42584

R:Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.

J. Gen. Virol. 79, 1197-1203, 1998

A:Title: The DNA sequence of equine herpesvirus-4.

A:Reference number: Z22173; MUID:98264497

A:Accession: T42584

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-529 <TEL>

A:Cross-references: EMBL:AF030027; NID:92605950; PIDN:AC59558.1; PID:92605966

A:Experimental source: strain NS80567

C:Genetics:

A:Gene: 40

C:Superfamily: varicella-zoster virus gene 38 protein

Query Match 81.0%; Score 34; DB 2; Length 529;

Best Local Similarity 87.5%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVNLLAEF 9

DB 423 EANLLAEF 430

RESULT 2

B96999

methy1-accepting chemotaxis protein [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: B96999

R:No1ling, J.; Bretton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: B96999

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-668 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK78781.1; PID:915023694; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC0805

Query Match 78.6%; Score 33; DB 2; Length 668;

Best Local Similarity 75.0%; Pred. No. 39;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SEVNLAE 8
||:||||
Db 71 SEINLAE 78

RESULT 3
ORHUL

LDL receptor precursor - human
C:Species: Homo sapiens (man)
C>Date: 17-May-1995 #sequence_revision 17-May-1985 #text_change 22-Jun-1999
C:Accession: A01383; A44176; A46751; S66642
C:Yamamoto, T.; Davis, C.G.; Brown, M.S.; Schneider, W.J.; Casey, M.L.; Goldstein, J.L.; Cell 39, 27-38, 1984
A>Title: The human LDL receptor: a cysteine-rich protein with Alu sequences in it its mR
A:Reference number: A01383; MUID:85024898
A:Accession: A01383
A:Molecule type: mRNA
A:Residues: 1-860 <YAM>
A:Cross-references: GB:L00352; GB:K02573; NID:g187094; PIDN:AAA56833.1; PID:g307121
R:Snudhof, T.C.; Goldstein, J.L.; Brown, M.S.; Russell, D.W.
Science 228, 815-822, 1985
A>Title: The LDL receptor gene: a mosaic of exons shared with different proteins.
A:Reference number: A44176; MUID:85218750
A:Accession: A44176
A:Molecule type: DNA
A>Status: nucleic acid sequence not shown
A:Residues: 1-184;193-314 <SUE>
A:Cross-references: GB:M10664
A>Note: translation of the nucleotide sequence is not complete
R:Russell, D.W.
submitted to GenBank, March 1985
A:Reference number: A46751
A:Accession: A46751
A:Molecule type: DNA
A:Residues: 1-860 <RUS>
A:Cross-references: GB:L00352; GB:K02573; NID:g187094; PIDN:AAA56833.1; PID:g307121
J:Yokode, M.; Pathak, R.K.; Hammer, R.E.; Brown, M.S.; Goldstein, J.L.; Anderson, R.G.W.
J. Cell Biol. 117, 39-46, 1992
A>Title: Cytoplasmic sequence required for basolateral targeting of LDL receptor in live
A:Reference number: A44535; MUID:92210646
A:Contents: annotation: basolateral targeting signal
R:Chen, W.J.; Goldstein, J.L.; Brown, M.S.
J. Biol. Chem. 265, 3116-3123, 1990
A>Title: NPXY, a sequence often found in cytoplasmic tails, is required for coated pit-
A:Reference number: A44541; MUID:90153958
A:Contents: annotation: internalization signal
R:Bieri, S.; Djordjevic, J.T.; Jamshidi, N.; Smith, R.; Kroon, P.A.
FEBS Lett. 371, 341-344, 1995
A>Title: Expression and disulfide-bond connectivity of the second ligand-binding repeat
A:Reference number: S66642; MUID:96013166
A:Contents: annotation: disulfide bonds
C:Comment: This transmembrane glycoprotein binds LDL, the major cholesterol-carrying lip
r-ligand complexes must first cluster into clathrin-coated pits.
C:Comment: An intrastand recombination event between two Alu sequences in the 3' untran
and cytoplasmic domains. Most of the receptors produced are secreted, but those that ad
grand complexes are not internalized.
C:Genetics:
A:Gene: GDB:LDLR
A:Cross-references: GDB:119362; OMIM:143890
A:Map position: 19p13.2-19p13.2
A:Introns: 23/1: 64/1: 105/1: 23/1: 273/1: 314/1: 354/1: 396/1: 453/2: 529/2: 569/1: 61
C:Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;
C:Keywords: cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipid
F:1-21/Domain: signal sequence #status predicted <Sig>
F:22-60/Product: LDL receptor #status predicted <EXT>
F:22-788/Domain: extracellular #status predicted <EXT>
F:27-63/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:68-104/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:109-143/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:148-184/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:197-231/Domain: LDL receptor ligand-binding repeat homology <LDL5>

F:236-270/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:276-313/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:318-352/Domain: EGF homology <EG1>
F:358-392/Domain: EGF homology <EG2>
F:399-438/Domain: LDL receptor WTD-containing repeat homology <YW1>
F:439-485/Domain: LDL receptor WTD-containing repeat homology <YW2>
F:486-528/Domain: LDL receptor WTD-containing repeat homology <YW3>
F:529-572/Domain: LDL receptor WTD-containing repeat homology <YW4>
F:573-615/Domain: LDL receptor WTD-containing repeat homology <YW5>
F:616-658/Domain: LDL receptor WTD-containing repeat homology <YW6>
F:667-711/Domain: EGF homology <EG3>
F:721-768/Region: clustered O-linked oligosaccharides
F:789-810/Domain: transmembrane #status predicted <TM>
F:811-860/Domain: intracellular #status predicted <INT>
F:825-829/Region: coated-pit mediated internalization signal
F:837-849/Region: basolateral targeting signal
F:862-75-95,89-104/Disulfide bonds: #status experimental
F:97,156,272,515,657/Binding site: carbonylate (asn) (covariant) #status predicted
F:318-329,325-338,340-352,358-368,364-377,379-392,667-681,677-696,698-711/Disulfide b

Query Match 78.6%; Score 33; DB 1; Length 860;
Best Local Similarity 87.5%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SEVNLAE 8
||:|||||
Db 637 SDVNLAE 644

RESULT 4
G72351
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: G72351
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A>Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316
A:Accession: G72351
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-132 <ARN>
A:Cross-references: GB:AE001738; GB:AE000512; NID:g4981158; PIDN:AD35732.1; PID:g498
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0648

Query Match 76.2%; Score 32; DB 2; Length 132;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAE 9
||:|||||
Db 86 SEINLAE 94

RESULT 5
T10822
1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14) - kidney bean (fragment)
N:Alternate names: ACC synthase
C:Species: Phaseolus vulgaris (kidney bean)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 04-Feb-2000
C:Accession: T10822
R:Pierson, C.M.; Reid, D.M.
submitted to the EMBL Data Library, March 1998
A:Description: The role of light in controlling ACC synthase activity and gene expres
A:Reference number: Z17175
A:Accession: T10822
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-279 <PID>
 A:Cross-references: EMBL:AF053355; NID:g2995852; PID:g2995853
 A:Experimental source: cultivar Taylor; leaf
 C:Genetics:
 A:Note: ACS1
 A:Note: Intron positions not resolved (incomplete sequence)
 C:Function:
 A:Description: catalyzes the formation of L-aminocyclopropane-1-carboxylic acid from S-adenosyl-L-methionine
 A:Pathway: ethylene biosynthesis
 C:Superfamily: 1-aminocyclopropane-1-carboxylate synthase
 C:Keywords: carbon-sulfur lyase; ethylene biosynthesis; phosphoprotein; pyridoxal phosphate
 F:190/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match
 Best Local Similarity 76.2%; Score 32; DB 2; Length 279;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
 |||||:|
 Db 129 SEVNLVDF 137

RESULT 6
 T09120
 basic leucine zipper protein - spinach
 C:Species: Spinacia oleracea (spinach)
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 20-Jun-2000
 C:Accession: T09120
 R:Bolle, C.; Lueberstedt, T.; Herranen, M.; Herrmann, R.; Oelmüller, R.
 submitted to the EMBL Data Library, January 1998
 A:Description: Molecular characterization of the spinach G-box binding protein family.
 A:Reference number: Z16572
 A:Accession: T09120
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-422 <BOL>
 A:Cross-references: EMBL:AJ223624
 C:Genetics:
 A:Gene: bZIP
 C:Superfamily: fava bean G-box-binding protein; fos/jun DNA-binding domain homology

Query Match
 Best Local Similarity 76.2%; Score 32; DB 2; Length 422;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 8
 |||||:|
 Db 324 SEVNLVDF 331

RESULT 7
 T10854
 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14) 7 - mung bean
 N:Alternate names: ACC synthase
 C:Species: Vigna radiata (mung bean)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 04-Feb-2000
 C:Accession: T10854
 R:Kim, W.
 submitted to the EMBL Data Library, August 1995
 A:Reference number: Z17182
 A:Accession: T10854
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-467 <KIM>
 A:Cross-references: EMBL:U34987; NID:g1006806; PID:g1006807
 A:Experimental source: tissue-type etiolated hypocotyl
 C:Genetics:
 A:Gene: ACS7
 C:Function:
 A:Description: catalyzes the rate-limiting first step of ethylene biosynthesis, the formation of S-adenosyl-L-methionine
 C:Superfamily: 1-aminocyclopropane-1-carboxylate synthase
 C:Keywords: carbon-sulfur lyase; ethylene biosynthesis; phosphoprotein; pyridoxal phosphate

F:272/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match
 Best Local Similarity 76.2%; Score 32; DB 2; Length 467;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
 |||||:|
 Db 211 SEVNLVDF 219

RESULT 8
 T10889
 1-aminocyclopropane-1-carboxylate synthase (EC 4.4.1.14) 6 - mung bean
 N:Alternate names: ACC synthase; S-adenosyl-L-methionine methyltransferase
 C:Species: Vigna radiata (mung bean)
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C:Accession: T10889; T10842
 R:Yoon, I.; Mori, H.; Kim, J.; Kang, B.; Imaseki, H.
 Plant Cell Physiol. 38, 217-224, 1997
 A:Title: VR-ACS6 is an auxin-inducible 1-aminocyclopropane-1-carboxylate synthase
 A:Reference number: Z17198; MUID:97294927
 A:Accession: T10889
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-472 <YOO>
 A:Cross-references: EMBL:AB000679; NID:g1813330; PID:BAAL9161.1; PID:g1813331
 R:Kim, W.
 submitted to the EMBL Data Library, August 1995
 A:Reference number: Z17182
 A:Accession: T10842
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-37, 'Q', '39-124', 'E', '126-389', 'OE', '392-425', 'NG', '428-442', 'T', '444-472 <KIM>
 A:Cross-references: EMBL:U34986; NID:g1006804; PID:g1006805
 A:Experimental source: tissue-type etiolated hypocotyls
 C:Genetics:
 A:Gene: ACS6
 C:Function:
 A:Description: catalyzes the conversion of S-adenosylmethionine to L-aminocyclopropane-1-carboxylate
 A:Pathway: ethylene biosynthesis
 C:Superfamily: 1-aminocyclopropane-1-carboxylate synthase
 C:Keywords: carbon-sulfur lyase; ethylene biosynthesis; phosphoprotein; pyridoxal phosphate
 F:275/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match
 Best Local Similarity 76.2%; Score 32; DB 1; Length 472;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
 |||||:|
 Db 211 SEVNLVDF 219

RESULT 9
 D95318
 NIFE oxidoreductase nife [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid
 C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: D95318
 R:Barnevelt, M.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B.
 ..; Kallman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yen, K.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: D95318
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-476 <KUR>
 A:Cross-references: GB:AE006469; PID:AAK65110.1; PID:g14523548; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Gilbert, F.; Finn, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.;
 Pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.

L.: Hyman, R.M.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebaull, P.; Vandemol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: nifE
 A:Genome: Plasmid
 C:Superfamily: dinitrogenase alpha chain; nitrogenase vanadium-iron protein alpha chain

Query Match 76.2%; Score 32; DB 2; Length 476;
 Best Local Similarity 55.6%; Pred. No. 44;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAEF 9
 Db 195 SDINLGEF 203

RESULT 10

E82333
 oligoribonuclease VC0341 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: E82333
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A:Reference number: A82035; PMID:20406833
 A:Accession: E82333
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-181 <HE>
 A:Cross-references: GB:AE004123; GB:AE003852; NID:99654756; PIDN:AAF93514.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC0341
 A:Map position: 1
 C:Superfamily: human conserved hypothetical protein DKF2p566E144.1

Query Match 73.8%; Score 31; DB 2; Length 181;
 Best Local Similarity 75.0%; Pred. No. 25;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 SEVNLLAE 8
 Db 35 SEVNLLAE 42

RESULT 11

T17795
 hypothetical protein A298L - Chlorella virus PCV-1
 C:Species: Chlorella virus PCV-1
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T17795
 R:Graves, M.V.; Van Etten, J.L.
 submitted to the EMBL Data Library, May 1999
 A:Reference number: Z18806
 A:Accession: T17795
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-225 <GRA>
 A:Cross-references: EMBL:U42580; NID:94028896; PIDN:AA09666.1
 A:Experimental source: specific host Chlorella strain NC64A
 C:Genetics:
 A:Note: A298L

Query Match 73.8%; Score 31; DB 2; Length 225;
 Best Local Similarity 85.7%; Pred. No. 32;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 VNLLAEF 9
 Db 9 VNLLAEF 15

RESULT 12

AD2820
 conserved hypothetical protein Atu1983 [imported] - Agrobacterium tumefaciens (strain
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: AD2820
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
 erge, G.; Gillet, W.; Grant, C.; Gentner, D.; Kutyavyn, T.; Levy, R.; Li, M.; McCl
 : Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AD2820
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-236 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AAL42978.1; PID:g17740438; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu1983
 A:Map position: circular chromosome

Query Match 73.8%; Score 31; DB 2; Length 236;
 Best Local Similarity 66.7%; Pred. No. 34;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAEF 9
 Db 137 SQVNLLASY 145

RESULT 13

D97598
 hypothetical protein AGR_C_3608 [imported] - Agrobacterium tumefaciens (strain C58, C
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
 C:Accession: D97598
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium*
 A:Reference number: A97359; PMID:11743194
 A:Accession: D97598
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-240 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK8741.1; PID:15157106; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_3608
 A:Map position: circular chromosome

Query Match 73.8%; Score 31; DB 2; Length 240;
 Best Local Similarity 66.7%; Pred. No. 34;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAEF 9
 Db 141 SQVNLLASY 149

RESULT 14

H87588
 hypothetical protein CC2743 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: H87588
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Hart, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: AB7249; MUID:21173698; PMID:11259647
 A:Accession: H87588
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-394 <STO>
 A:Cross-references: GB:AE005673; NID:g13424336; PIDN:AAK24708.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC2743

Query Match 73.8%; Score 31; DB 2; Length 394;
 Best Local Similarity 87.5%; Pred. No. 60;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAE 8
 |||||
 Db 370 SEVNLXAE 377

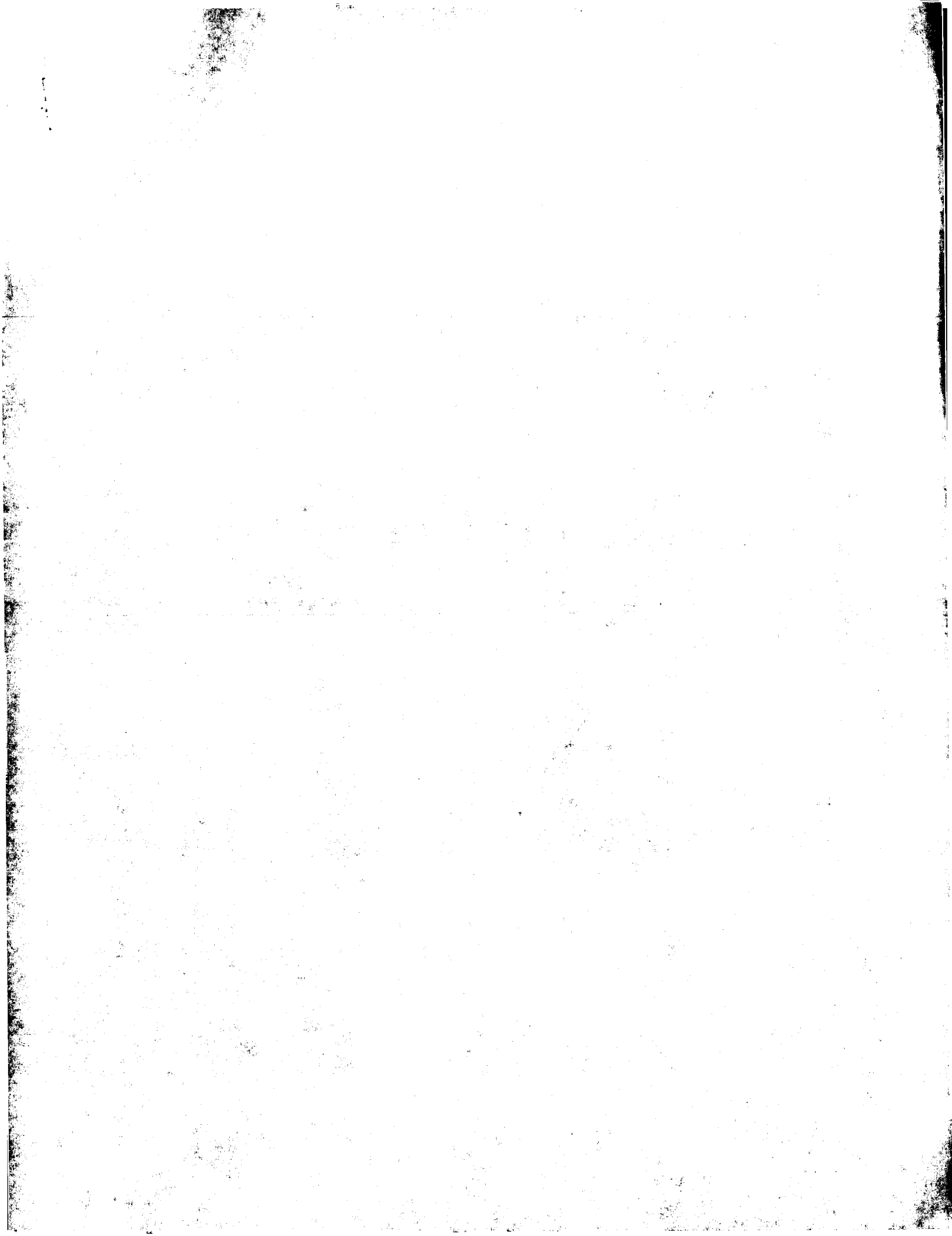
RESULT 15

T15020
 probable porphyrin biosynthetic enzyme - *Yersinia pestis* plasmid pMT1
 C:Species: *Yersinia pestis*
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Sep-1999
 C:Accession: T15020; T14700
 R:Jandler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.
 Infect. Immun. 66, 5731-5742, 1998
 A:Title: Complete DNA sequence and detailed analysis of the *Yersinia pestis* K1M5 plasmid
 A:Reference number: Z18268; MUID:99043898
 A:Accession: T15020
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-411 <LIN>
 A:Cross-references: EMBL:AF074611; NID:g3883003; PID:g3883103; PIDN:AAC82763.1
 R:Hu, P.; Elliott, J.; McCreedy, P.; Skowronski, E.; Garnea, J.; Kobayashi, A.; Carrano,
 submitted to the EMBL Data Library, March 1998
 A:Description: Structural organization of virulence determinants in three *Yersinia pestis*
 A:Reference number: Z18168
 A:Accession: T14700
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-411 <RUP>
 A:Cross-references: EMBL:AF053947; NID:g2996286; PID:g2996336; PIDN:AAC13216.1
 C:Genetics:
 A:Gene: COB5
 A:Genome: plasmid pMT1

Query Match 73.8%; Score 31; DB 2; Length 411;
 Best Local Similarity 75.0%; Pred. No. 63;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SEVNLAE 8
 :|||:|
 Db 326 AEVNLAE 333

Search completed: October 30, 2002, 12:31:37
 Job time : 4.72482 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:24:20 ; Search time 1.28256 Seconds
(without alignments)
171.400 Million cell updates/sec

Title: US-09-724-571-91

Perfect score: 42

Sequence: 1 SEVNLLAEF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest1.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	81.0	9	3	US-08-802-981-219
2	34	81.0	10	2	US-08-659-984A-19
3	34	81.0	10	4	US-08-660-531-19
4	34	81.0	11	5	PCT-US94-07043A-3
5	34	81.0	21	2	US-08-659-984A-18
6	34	81.0	21	3	US-08-802-981-112
7	34	81.0	21	4	US-08-660-531-18
8	34	81.0	30	2	US-08-659-984A-17
9	34	81.0	30	4	US-08-660-531-17
10	34	81.0	33	2	US-08-659-984A-16
11	34	81.0	33	4	US-08-660-531-16
12	34	81.0	42	2	US-08-659-984A-15
13	34	81.0	42	4	US-08-660-531-15
14	34	81.0	506	2	US-08-659-984A-21
15	34	81.0	506	4	US-08-660-531-21
16	34	81.0	506	4	US-09-054-334-4
17	33	78.6	280	4	US-08-652-877-6
18	33	78.6	860	1	US-08-476-515A-6
19	33	78.6	860	1	US-08-092-817-4
20	33	78.6	1410	2	US-08-470-058-4
21	33	78.6	1410	3	US-09-037-188-4
22	33	78.6	1410	4	US-09-285-310-4
23	32	76.2	942	4	US-09-171-461-12
24	30	71.4	40	4	US-07-861-458C-118
25	30	71.4	67	6	5196333-10
26	30	71.4	103	3	US-08-339-708A-12
27	30	71.4	493	6	5196333-4

28	30	71.4	564	4	US-09-360-197-16	Sequence 16, Appl
29	30	71.4	753	4	US-07-861-458C-98	Sequence 98, Appl
30	30	71.4	755	4	US-07-861-458C-99	Sequence 99, Appl
31	29	69.0	27	1	US-08-141-324-12	Sequence 12, Appl
32	29	69.0	27	1	US-08-541-902-12	Sequence 12, Appl
33	29	69.0	27	1	US-08-942-012B-33	Sequence 33, Appl
34	28	66.7	9	3	US-08-802-981-223	Sequence 220, App
35	28	66.7	9	3	US-08-802-981-220	Sequence 223, App
36	28	66.7	21	3	US-08-802-981-113	Sequence 113, App
37	28	66.7	21	3	US-08-802-981-116	Sequence 116, App
38	28	66.7	74	1	US-08-543-238-5	Sequence 5, Appl
39	28	66.7	74	1	US-08-420-526-5	Sequence 5, Appl
40	28	66.7	237	2	US-08-933-750C-32	Sequence 32, Appl
41	28	66.7	237	4	US-09-234-613-32	Sequence 32, Appl
42	27	64.3	12	4	US-09-399-494-22	Sequence 22, Appl
43	27	64.3	31	4	US-09-399-494-11	Sequence 11, Appl
44	27	64.3	34	3	US-09-044-536A-21	Sequence 21, Appl
45	27	64.3	64	2	US-08-687-559-3	Sequence 3, Appl

ALIGNMENTS

```
RESULT 1
US-08-802-981-219
; Sequence 219, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,981
; FILING DATE: 20-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 016865-000300US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-802-981-219
```

Query Match 81.0%; Score 34; DB 3; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.7e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SEVNLLAEF 9
|||||
Db 1 SEVNLLAEF 9

```
RESULT 2
US-08-659-984A-19
; Sequence 19, Application US/08659984A
; Patent No. 5942400
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Sinha, Sukanto
; APPLICANT: Jacobson-Croak, Kirsten L.
; TITLE OF INVENTION: Assays for Detecting Beta-Secretase
; TITLE OF INVENTION: Inhibition
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Ctr., 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,984A
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,152
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-0028100S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: one-of(1)
; OTHER INFORMATION: /note= "N-terminal Ser is acetylated."
US-08-659-984A-19

Query Match      81.0%; Score 34; DB 2; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.33;
Matches      8; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

QY      1 SEVNLDAEF 9
Db      1 SEVNLDAEF 9

RESULT 3
US-08-660-531-19
; Sequence 19, Application US/08660531
; Patent No. 6221645
; GENERAL INFORMATION:
; APPLICANT: Chrysler, Susanna M.S.
; APPLICANT: Sinha, Sukanto
; APPLICANT: Keim, Pamela S.
; APPLICANT: Anderson, John P.
; TITLE OF INVENTION: Beta-Secretase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
```

```
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,531
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,498
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-0022100S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: one-of(1)
; OTHER INFORMATION: /note= "N-terminal Ser is acetylated."
US-08-660-531-19

Query Match      81.0%; Score 34; DB 4; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.33;
Matches      8; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

QY      1 SEVNLDAEF 9
Db      1 SEVNLDAEF 9

RESULT 4
PCT-US94-07043A-3
; Sequence 3, Application PC/TUS9407043A
; GENERAL INFORMATION:
; APPLICANT: Tamburini, Paul P.; Benz, G nter; H bich,
; APPLICANT: Dieter; Dreyer, Robert N.; Koenig, Gerhard
; TITLE OF INVENTION: CATHEPSIN D IS AN AMYLOIDGENIC
; TITLE OF INVENTION: PROTEASE IN ALZHEIMER S DISEASE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Miles Inc.
; STREET: 400 Morgan Lane
; CITY: West Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06516
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage
; COMPUTER: Sharp PC 4600
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07043A
; FILING DATE: June 21, 1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10889
```

;; FILING DATE: November 12, 1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/995,660
;; FILING DATE: December 16, 1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/880,914
;; FILING DATE: May 11, 1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Pamela A. Simonton
;; REGISTRATION NUMBER: 31,060
;; REFERENCE/DOCKET NUMBER: MTI 224.3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (203) 937-2340
;; TELEFAX: (203) 937-2795
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
PCT-US94-07043A-3

Query Match 81.0%; Score 34; DB 5; Length 11;
Best Local Similarity 88.9%; Pred. No. 0.36;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAEF 9
||| |||
DB 2 SEVNLLAEF 10

RESULT 5
US-08-659-984A-18
; Sequence 18, Application US/08659984A
; Patent No. 5942400
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Simha, Sukanto
; APPLICANT: Jacobson-Ciroak, Kirsten L.
; TITLE OF INVENTION: Assays for Detecting Beta-Secretase
; TITLE OF INVENTION: Inhibition
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Ctr., 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,984A
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,152
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-002810US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-08-659-984A-18

Query Match 81.0%; Score 34; DB 2; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.74;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAEF 9
||| |||
DB 1 SEVNLLAEF 9

RESULT 6
US-08-802-981-112
; Sequence 112, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,981
; FILING DATE: 20-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 016865-000300US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product="Aib"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /product="Acp"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 16
; OTHER INFORMATION: /product="Acp"
US-08-802-981-112

Query Match 81.0%; Score 34; DB 3; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.74;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAEF 9
||| |||
DB 6 SEVNLLAEF 14

RESULT 7
US-08-660-531-18
Sequence 18, Application US/08660531
Patent No. 6221645
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Keim, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,531
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,498
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-660-531-18

Query Match 81.0%; Score 34; DB 4; Length 21;
Best Local Similarity 88.9%; Pred. No. 0.74;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNDAER 9
DB 1 SEVNDAER 9

RESULT 8
US-08-659-984A-17
Sequence 17, Application US/08659984A
Patent No. 5942400
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Sinha, Sukanto
APPLICANT: Jacobson-Croak, Kirsten L.
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California

COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,984A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-659-984A-17

Query Match 81.0%; Score 34; DB 2; Length 30;
Best Local Similarity 88.9%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNDAER 9
DB 22 SEVNDAER 30

RESULT 9
US-08-660-531-17
Sequence 17, Application US/08660531
Patent No. 6221645
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Keim, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,531
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,498
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-660-531-17

Query Match 81.0%; Score 34; DB 4; Length 30;
Best Local Similarity 88.9%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
|||||
Db 22 SEVNDAEF 30

RESULT 10
US-08-659-984A-16
Sequence 16, Application US/08659984A
Patent No. 5942400
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Sinha, Sukanto
APPLICANT: Jacobson-Croak, Kirsten L.
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,984A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-659-984A-16

Query Match 81.0%; Score 34; DB 2; Length 33;
Best Local Similarity 88.9%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 SEVNLAEF 9
|||||

Db 13 SEVNDAEF 21

RESULT 11
US-08-660-531-16
Sequence 16, Application US/08660531
Patent No. 6221645
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Keim, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,531
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,498
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-660-531-16

Query Match 81.0%; Score 34; DB 4; Length 33;
Best Local Similarity 88.9%; Pred. No. 1.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
|||||
Db 13 SEVNDAEF 21

RESULT 12
US-08-659-984A-15
Sequence 15, Application US/08659984A
Patent No. 5942400
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Sinha, Sukanto
APPLICANT: Jacobson-Croak, Kirsten L.
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco

STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,984A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-659-984A-15

Query Match 81.0%; Score 34; DB 2; Length 42;
Best Local Similarity 88.9%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAEE 9
|||||
Db 22 SEVNLAEE 30

RESULT 13
US-08-660-531-15
Sequence 15, Application US/08660531
Patent No. 6221645
GENERAL INFORMATION:
APPLICANT: Chrysler, Susanna M.S.
APPLICANT: Sinha, Sukanto
APPLICANT: Kelm, Pamela S.
APPLICANT: Anderson, John P.
TITLE OF INVENTION: Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/660,531
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,498
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541

REFERENCE/DOCKET NUMBER: 15270-002210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-660-531-15

Query Match 81.0%; Score 34; DB 4; Length 42;
Best Local Similarity 88.9%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAEE 9
|||||
Db 22 SEVNLAEE 30

RESULT 14
US-08-659-984A-21
Sequence 21, Application US/08659984A
Patent No. 5942400
GENERAL INFORMATION:
APPLICANT: Anderson, John P.
APPLICANT: Sinha, Sukanto
APPLICANT: Jacobson-Croak, Kirsten L.
TITLE OF INVENTION: Assays for Detecting Beta-Secretase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Ctr., 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,984A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,152
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-002810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-984A-21

Query Match 81.0%; Score 34; DB 2; Length 506;
Best Local Similarity 88.9%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAEE 9
|||||

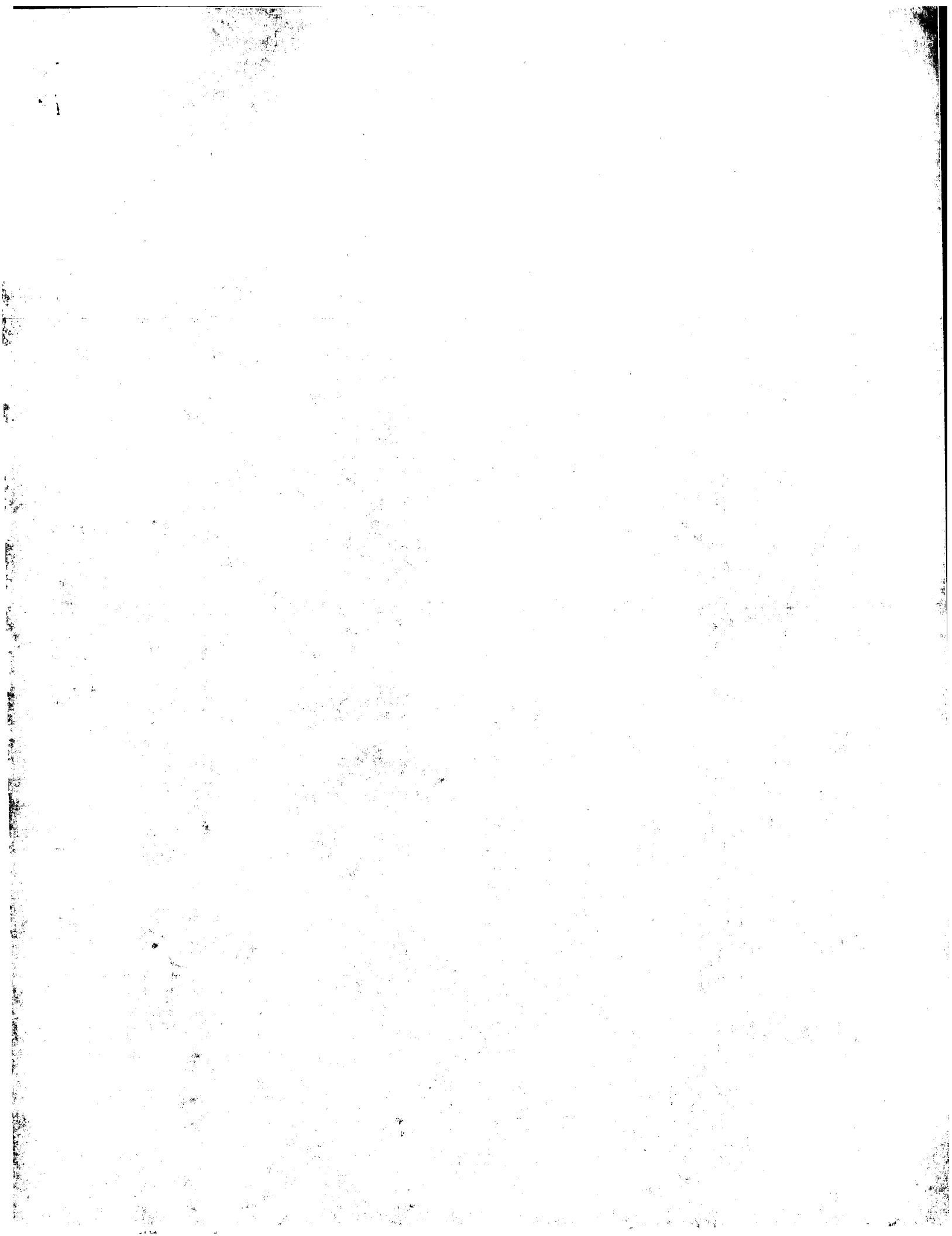
Db 403 SEVNLDAEF 411

RESULT 15
US-08-660-531-21
; Sequence 21, Application US/08660531
; Patent No. 6221645
; GENERAL INFORMATION:
; APPLICANT: Chrysler, Susanna M.S.
; APPLICANT: Sinha, Sukanto
; APPLICANT: Keim, Pamela S.
; APPLICANT: Anderson, John P.
; TITLE OF INVENTION: Beta-Secretase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Ctr., 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,531
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,498
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-002210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-660-531-21

Query Match 81.0%; Score 34; DB 4; Length 506;
Best Local Similarity 88.9%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLDAEF 9
|||||
Db 403 SEVNLDAEF 411

Search completed: October 30, 2002, 12:32:35
Job time : 2.28256 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:21:30 ; Search time 2.49631 Seconds
(without alignments)
554.401 Million cell updates/sec

Title: US-09-724-571-81

Perfect score: 34
Sequence: 1 EVMXVAEF 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17294929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB	ID	Description
1	29	85.3	195	16	097F65	097F65 clostridium
2	28	82.4	776	10	0941L7	0941L7 populus tire
3	28	82.4	3414	12	091C40	091C40 langat viru
4	28	82.4	3414	12	091G39	091G39 langat viru
5	27	79.4	184	17	026364	026364 methanother
6	27	79.4	201	16	034914	034914 bacillus su
7	27	79.4	338	17	09HL69	09HL69 thermoplasm
8	27	79.4	598	17	058774	058774 pyrococcus
9	27	79.4	601	16	092U63	092U63 thizobium m
10	27	79.4	775	10	09FE38	09FE38 arabidopsis
11	27	79.4	802	10	09SUR2	09SUR2 arabidopsis
12	27	79.4	1158	5	096554	096554 caenorhabdi
13	27	79.4	1560	5	026644	026644 strongyloce
14	26	76.5	46	16	092ZL0	092ZL0 rhizobium m
15	26	76.5	92	9	09FZT5	09FZT5 pseudomonas
16	26	76.5	150	2	052322	052322 escherichia

17	26	76.5	194	10	0946P3	0946P3 uncultured
18	26	76.5	252	16	0927W5	0927W5 chlamydia p
19	26	76.5	278	10	09FXF4	09FXF4 arabidopsis
20	26	76.5	294	10	09M2H4	09M2H4 arabidopsis
21	26	76.5	324	10	092P05	092P05 arabidopsis
22	26	76.5	363	4	09H266	09H266 homo sapien
23	26	76.5	490	16	092T24	092T24 rhizobium m
24	26	76.5	498	16	09KN11	09KN11 vibrio chol
25	26	76.5	533	16	092BS4	092BS4 listeria in
26	26	76.5	543	16	09KD12	09KD12 bacillus ha
27	26	76.5	559	1	09UWN9	09UWN9 sulfolobus
28	26	76.5	559	1	055088	055088 sulfolobus
29	26	76.5	561	17	095867	095867 sulfolobus
30	26	76.5	562	17	09UZ10	09UZ10 pyrococcus
31	26	76.5	579	2	09KZD3	09KZD3 streptomyc
32	26	76.5	587	16	09UZP8	09UZP8 neisseria m
33	26	76.5	587	16	09JUT3	09JUT3 neisseria m
34	26	76.5	623	16	09PAA1	09PAA1 xylella fas
35	26	76.5	642	3	09C260	09C260 neurospora
36	26	76.5	778	11	099MP4	099MP4 mus musculu
37	26	76.5	791	4	09P2E2	09P2E2 homo sapien
38	26	76.5	821	2	09XBW4	09XBW4 porphyromon
39	26	76.5	1014	16	09K9C6	09K9C6 bacillus ha
40	26	76.5	1085	5	001585	001585 caenorhabdi
41	26	76.5	1249	5	097043	097043 drosophila
42	26	76.5	1249	5	09VF33	09VF33 drosophila
43	25	73.5	127	12	091BF3	091BF3 spodiopetra
44	25	73.5	136	11	063172	063172 rattus norv
45	25	73.5	137	4	09N248	09N248 homo sapien

ALIGNMENTS

RESULT 1

097F65 ID 097F65 PRELIMINARY: PRT: 195 AA.

AC 097F65: 01-OCT-2001 (TREMBLrel. 18, Created)

DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE THYMIDINE KINASE.

GN CAC2887.

OS Clostridium acetobutylicum.

OC Bacteria: Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

CC Clostridium.

OX NCBI_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=11466286;

RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,

RA Gibson R., Lee H.M., Dubois J., Qiu D., Hiltl J., Wolf Y.I.,

RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,

RA Bennett G.N., Koonin E.V., Smith D.R.;

RT *Genome sequence and comparative analysis of the solvent-producing

RT bacterium Clostridium acetobutylicum";

RL J. Bacteriol. 183:4823-4838(2001).

DR EMBL: AE007786; AAK80830.1; -.

DR InterPro: IPR000345; CytC_heme_bind.

DR InterPro: IPR001267; Thymd_kin_cell.

DR Pfam: PF00265; TK; 1.

DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN_1.

DR PROSITE: PS00603; TK_CELLULAR_TYPE; UNKNOWN_1.

KW kinase; Complete proteome.

SO SEQUENCE 195 AA; 22303 MW; 6EDB551986B203BA CRC64;

Query Match 85.3%; Score 29; DB 16; Length 195;

Best Local Similarity 62.5%; Pred. No. 18;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVMXVAEF 8

|||||

Db 130 ELMATAEF 137

RESULT 2

0941L7 PRELIMINARY; PRT; 776 AA.

AC 0941L7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE HIGH-AFFINITY POTASSIUM UPTAKE TRANSPORTER.
 GN KUP1.
 OS Populus tremula x Populus tremuloides.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids I; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=47664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Langer K., Ache P., Fromm J., Hedrich R.;
 RT "PKUP contributes to wood formation";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ299422; CAC39168.1; -
 SQ SEQUENCE 776 AA; 87303 MW; BEC03D57ED0869BC CRC64;

Query Match 82.4%; Score 28; DB 10; Length 776;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 VMXVAEF 8
 |||||
 Db 604 VMSVAEF 610

RESULT 3

091G40 PRELIMINARY; PRT; 3414 AA.

AC 091G40;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE POLYPROTEIN PRECURSOR.
 OS Langat virus (strain TP21).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 OX NCBI_TaxID=31638;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TP21;
 RX MEDLINE=20192178; PubMed=10725214;
 RA Campbell M.S., Pietnev A.G.;
 RT "Infectious cDNA clones of Langat tick-borne flavivirus that differ
 from their parent in peripheral neurovirulence.";
 RL Virology 269:225-237(2000).
 DR EMBL; AF253419; AAF75259.1; -
 DR HSSP; P14336; 1SVB.
 DR MEROPS; S07.001; -
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR000336; Flavi_glycoprote.
 DR InterPro; IPR001850; Flavi_helicase.
 DR InterPro; IPR001157; Flavi_NSI.
 DR InterPro; IPR001528; Flavi_NSA4.
 DR InterPro; IPR000208; Flavi_NS5.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00869; Flavi_glycoprote; 1.
 DR Pfam; PF00949; Flavi_helicase; 1.
 DR Pfam; PF00948; Flavi_NSI; 1.
 DR Pfam; PF01349; Flavi_NSA4; 1.
 DR Pfam; PF00972; Flavi_NS5; 1.
 DR ProDom; PD001496; Flavi_NSI; 1.
 DR ProDom; PD001556; Flavi_glycoprote; 1.
 DR SMART; SM00490; HELICC; 1.
 KW Signal.

FT SIGNAL 97 116 POTENTIAL.
 FT CHAIN 1 96 CAPSID PROTEIN.
 FT CHAIN 117 280 MEMBRANE-ASSOCIATED GLYCOPROTEIN
 FT CHAIN 281 779 PRECURSOR
 FT CHAIN 780 1128 ENVELOPE MEMBRANE-ASSOCIATED
 FT CHAIN 1129 1358 GLYCOPROTEIN.
 FT CHAIN 1359 1489 NONSTRUCTURAL PROTEIN NS1.
 FT CHAIN 1490 2110 NONSTRUCTURAL PROTEIN NS2A.
 FT CHAIN 2111 2259 NONSTRUCTURAL PROTEIN NS2B.
 FT CHAIN 2260 2511 NONSTRUCTURAL PROTEIN NS3.
 FT CHAIN 2512 3414 NONSTRUCTURAL PROTEIN NS4.
 FT CHAIN 3414 378135 NONSTRUCTURAL PROTEIN NS4B.
 FT CHAIN 378135 8DE86A46A0E8F7E3 CRC64;
 SQ SEQUENCE 3414 AA; 378135 MW; 8DE86A46A0E8F7E3 CRC64;

Query Match 82.4%; Score 28; DB 12; Length 3414;
 Best Local Similarity 85.7%; Pred. No. 5.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 VMXVAEF 8
 |||||
 Db 929 VMTVAEF 935

RESULT 4

091G39 PRELIMINARY; PRT; 3414 AA.

AC 091G39;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE POLYPROTEIN PRECURSOR.
 OS Langat virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 OX NCBI_TaxID=11085;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATTENUATED STRAIN E5;
 RX MEDLINE=20192178; PubMed=10725214;
 RA Campbell M.S., Pietnev A.G.;
 RT "Infectious cDNA clones of Langat tick-borne flavivirus that differ
 from their parent in peripheral neurovirulence.";
 RL Virology 269:225-237(2000).
 DR EMBL; AF253420; AAF75260.1; -
 DR HSSP; P14336; 1SVB.
 DR MEROPS; S07.001; -
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR000336; Flavi_glycoprote.
 DR InterPro; IPR001850; Flavi_helicase.
 DR InterPro; IPR001157; Flavi_NSI.
 DR InterPro; IPR001528; Flavi_NSA4.
 DR InterPro; IPR000208; Flavi_NS5.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00869; Flavi_glycoprote; 1.
 DR Pfam; PF00949; Flavi_helicase; 1.
 DR Pfam; PF00948; Flavi_NSI; 1.
 DR Pfam; PF01349; Flavi_NSA4; 1.
 DR Pfam; PF00972; Flavi_NS5; 1.
 DR ProDom; PD001496; Flavi_NSI; 1.
 DR ProDom; PD001556; Flavi_glycoprote; 1.
 DR SMART; SM00490; HELICC; 1.
 KW SIGNAL.
 FT SIGNAL 97 116 POTENTIAL.
 FT CHAIN 1 96 CAPSID PROTEIN.
 FT CHAIN 117 280 MEMBRANE-ASSOCIATED GLYCOPROTEIN
 FT CHAIN 281 779 PRECURSOR.
 FT CHAIN 780 1128 ENVELOPE MEMBRANE-ASSOCIATED
 FT CHAIN 1129 1358 GLYCOPROTEIN.
 FT CHAIN 1359 1489 NONSTRUCTURAL PROTEIN NS1.
 FT CHAIN 1490 2110 NONSTRUCTURAL PROTEIN NS2A.
 FT CHAIN 2111 2259 NONSTRUCTURAL PROTEIN NS2B.
 FT CHAIN 2260 2511 NONSTRUCTURAL PROTEIN NS3.
 FT CHAIN 2512 3414 NONSTRUCTURAL PROTEIN NS4.
 FT CHAIN 3414 378135 NONSTRUCTURAL PROTEIN NS4B.
 FT CHAIN 378135 8DE86A46A0E8F7E3 CRC64;

```

FT CHAIN 2111 2259 NONSTRUCTURAL PROTEIN NS4A.
FT CHAIN 2260 2511 NONSTRUCTURAL PROTEIN NS4B.
FT CHAIN 2512 3414 NONSTRUCTURAL PROTEIN NS5.
SQ SEQUENCE 3414 AA; 378014 MW; BF8F8ACEAB96D534 CRC64;

Query Match
Best Local Similarity 82.4%; Score 28; DB 13; Length 3414;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 VMXVAEF 8
DB 929 VMXVAEF 935

RESULT 5
ID 026364 PRELIMINARY; PRT; 184 AA.
AC 026364;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE DNA-DEPENDENT RNA POLYMERASE, SUBUNIT E'.
GN MTH264.
OS Methanothermobacter thermautotrophicus.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RC MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Viore R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pletrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nollings J., Reeve J.N.;
RT *Complete genome sequence of Methanobacterium thermautotrophicum
RT delah: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7135(1997).
DR EMBL: AE000812; AAB84770.1; -
DR HSSP: P05055; ISRO.
DR InterPro: IPR003029; S1.
DR Pfam: PF00575; S1; 1.
DR SMART: SM00316; S1; 1.
KW Complete proteome.
SQ SEQUENCE 184 AA; 20695 MW; 37AAB7B39D631813 CRC64;

Query Match
Best Local Similarity 79.4%; Score 27; DB 17; Length 184;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVXVAEF 8
DB 88 EVIEIAEF 95

RESULT 6
ID 034914 PRELIMINARY; PRT; 201 AA.
AC 034914;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE YTHO.
GN YTHO.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=96312354; PubMed=8763940;
RA Bower S., Perkins J.B., Yocum R.R., Howitt C.L., Rahim P., Pero J.;
RT "Cloning, sequencing, and characterization of the Bacillus subtilis
RT biotin biosynthetic operon.";
RL J. Bacteriol. 178:4122-4130(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Lapidus A., Galleron N., Sorokin A., Ehrlich D.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RC MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriis R., Boutsier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Enlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.-J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kleerr-Blanchard M., Klein C.,
RA Kodayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Moore D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,
RA Parro V., Pohl T.M., Portetle D., Porwollik S., Prescott A.M.,
RA Priesecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadleir I.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
RA Seliguchi J., Sekowska A., Serron S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terstera P., Tognoni K.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT *The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RC Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF008220; AAC00267.1; -
DR EMBL: Z99119; CAB14996.1; -
KW Complete proteome.
SQ SEQUENCE 201 AA; 22733 MW; CBFAD9B6BA348F06 CRC64;

Query Match
Best Local Similarity 79.4%; Score 27; DB 16; Length 201;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVXVAEF 8
DB 22 DIMDIAEF 29

RESULT 7
ID 09HL69 PRELIMINARY; PRT; 338 AA.
AC 09HL69;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN TA0362.
GN TA0362.
OS Thermoplasma acidophilum.

```

OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
 OC Thermoplasma.
 OC NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermophilic scavenger Thermoplasma
 RT acidophilum.";
 RL Nature 407:508-513(2000).
 RL EMBL; AL45064; CAC11506.1; -
 DR InterPro: IPR001005; MYB_DNA_bind.
 DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 338 AA; 39348 MW; D86C955529E9A43E CRC64;

Query Match 79.4%; Score 27; DB 17; Length 338;
 Best Local Similarity 62.5%; Pred. No. 1e+02; 1; Indels 0; Gaps 0;
 Matches 5; Conservative 2; Mismatches 1;

OY 1 EVMXVAF 8
 DB 329 EIMPLAF 336

RESULT 8
 ID 058774 PRELIMINARY; PRT; 598 AA.

AC 058774;
 DT 01-AUG-1998 (TRENBLREL. 07, Created)
 DT 01-AUG-1998 (TRENBLREL. 07, Last sequence update)
 DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)
 DE HYPOTHEICAL 70.2 KDA PROTEIN PH1023.
 GN PH1023.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OC NCBI_TaxID=33953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.I., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 RL EMBL; AF000004; BAA30120.1; -
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 598 AA; 70188 MW; 3E1ACD0F8EDC7661 CRC64;

Query Match 79.4%; Score 27; DB 17; Length 598;
 Best Local Similarity 62.5%; Pred. No. 1.8e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVMXVAF 8
 DB 280 EIVVAF 287

RESULT 9
 ID 092063 PRELIMINARY; PRT; 601 AA.

AC 092063;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE PUTATIVE 1,4-ALPHA-GLOCAN BRANCHING ENZYME PROTEIN (EC

DE 2.4.1.18).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Plasmid pSymB (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OC NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA Golding B., Puehler A.;
 RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AL603646; CAC49676.1; -
 KW Transferase; Glycosyltransferase; Plasmid; Hypothetical protein;
 KM Complete proteome.
 SO SEQUENCE 601 AA; 66954 MW; 024DA322A7B72C2E CRC64;

Query Match 79.4%; Score 27; DB 16; Length 601;
 Best Local Similarity 62.5%; Pred. No. 1.8e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVMXVAF 8
 DB 151 EIMPVAF 158

RESULT 10
 ID 09FE38 PRELIMINARY; PRT; 775 AA.

AC 09FE38;
 DT 01-MAR-2001 (TRENBLREL. 16, Created)
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE TINY ROOT HAIR 1 PROTEIN.
 GN TRH1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rillas S., Debrosses G., Haralampidis K., Vicente-Agullo F.,
 RA Feldmann K.A., Grabov A., Dolan L., Hatzopoulos P.;
 RT "A Potassium Transporter Required for Tip Growth in Arabidopsis.";
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ296156; CAC16138.1; -
 DR EMBL; AJ296155; CAC16137.1; -
 DR InterPro: IPR003855; K+-transprtr.
 DR Pfam: PF02705; K_trans; 1.
 SO SEQUENCE 775 AA; 86842 MW; B0C5068B48E8180 CRC64;

Query Match 79.4%; Score 27; DB 10; Length 775;
 Best Local Similarity 71.4%; Pred. No. 2.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 VMXVAF 8
 DB 602 VMSVAF 608

RESULT 11
 ID 09SUR2 PRELIMINARY; PRT; 802 AA.

AC 09SUR2;
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE PUTATIVE POTASSIUM TRANSPORT PROTEIN.

GN F0016.110 OR ATG23640.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Medler H., Medler E., Wambutt R., Hohnsels J., Mewes H.W.,
RA Mayer K.F.X., Scheller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Medler H., Medler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL033394; CAA23030.1; -;
DR EMBL; AL161559; CAB79319.1; -;
DR InterPro: IPR003855; K+_transprtr.
DR Pfam: PF02705; K_trans; 1.
SO SEQUENCE 802 AA; 89817 MW; 617AF5F76B9BB60 CRC64;
Query Match 79.4%; Score 27; DB 10; Length 802;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 EVXVAEF 8
Db 629 VMSIAEF 635
RESULT 12
096554
ID 096554 PRELIMINARY; PRT; 1158 AA.
AC 096554;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN Y75B7AL.4.
GN Y75B7AL.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RC STRAIN-BRISTOL N2;
RA Lamar B.;
RT "The sequence of C. elegans cosmid Y75B7AL.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024669; AAK68606.1; -;
SO SEQUENCE 1158 AA; 127146 MW; 63D522E943FF01B CRC64;

Query Match 79.4%; Score 27; DB 5; Length 1158;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVXVAEF 8
Db 563 EOMVAEF 570
RESULT 13
026644
ID 026644 PRELIMINARY; PRT; 1560 AA.
AC 026644;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CALCIUM-BINDING PROTEIN.
GN END016.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinozoa; Euechinozoa; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RA Soltysik-Espanola M., Klinzing D.C., Pfarr K., Burke R.D., Ernst S.G.;
RT "Endo16, a large multi-domain protein found on the surface and ECM of
RT endodermal cells during sea urchin gastrulation, binds calcium.";
RL Dev. Biol. 0:0-0(1994).
DR EMBL; L34680; AAA30047.1; -;
DR InterPro: IPR001614; Myelin_PLP.
DR Pfam: PF002861; Reeler.
DR SMART: SM00002; PLP; 1.
SO SEQUENCE 1560 AA; 175228 MW; 2D1A0E0D8E97CE70 CRC64;
Query Match 79.4%; Score 27; DB 5; Length 1560;
Best Local Similarity 62.5%; Pred. No. 4.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVXVAEF 8
Db 705 ELLVAEF 712
RESULT 14
0922L0
ID 0922L0 PRELIMINARY; PRT; 46 AA.
AC 0922L0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN SMA0833.
GN SMA0833.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-1021;
RC MEDLINE-21396509; PubMed-11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Gilbert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kaiman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Teh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSyma megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL; AE007236; AAK65112.1; -;

KW plasmid: Hypothetical protein: Complete proteome.
 SQ SEQUENCE 46 AA; 5236 MW; 3AD9B653CB1B4813 CRC64;

Query Match 76.5%; Score 26; DB 16; Length 46;
 Best Local Similarity 71.4%; Pred. No. 26;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVMXVAE 7
 |::|::|
 Db 17 EIMXVAE 23

RESULT 15

09FZT5 PRELIMINARY; PRT; 92 AA.
 AC 09FZT5;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
 DE P3B.
 CN 3B.
 OS Pseudomonas bacteriophage phi-13.
 OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
 OX NCBI_TaxID=134554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99350412; PubMed=10419946;
 RA Mindich L., Qiao X., Qiao J., Onodera S., Romantschuk M.,
 Hoogstraaten D.;
 RT "Isolation of additional bacteriophages with genomes of segmented
 RT double-stranded RNA.";
 RL J. Bacteriol. 181:4505-4508(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20473938; PubMed=11017801;
 RA Qiao X., Qiao J., Onodera S., Mindich L.;
 RT "Characterization of phi13, a bacteriophage related to phi6 and
 RT containing three dsRNA genomic segments.";
 RL Virology 275:218-224(2000).
 DR EMBL; AF261667; AAG00441.1; -;
 SQ SEQUENCE 92 AA; 9904 MW; B52E88EDB8720740 CRC64;

Query Match 76.5%; Score 26; DB 9; Length 92;
 Best Local Similarity 50.0%; Pred. No. 51;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVMXVAEF 8
 |::|::|
 Db 47 EILSVSEF 54

Search completed: October 30, 2002, 12:30:10
 Job time : 5.49631 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:20:50 ; Search time 3.60442 Seconds
(without alignments)
277.344 Million cell updates/sec

Title: US-09-724-571-91

Sequence: 1 SEYNLLAEF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_032802.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	21	AA07882
2	40	95.2	19	22	AA07882
3	39	92.9	14	21	AA07888
4	38	90.5	9	21	AA07887
5	37	88.1	9	21	AA07879
6	36	85.7	9	21	AA07883
7	35	83.3	649	22	AB04111
8	34	81.0	9	19	AA082081
9	34	81.0	9	21	AA07874
10	34	81.0	9	21	AA07881
11	34	81.0	9	21	AA07894

12	34	81.0	9	22	AA073297	Protease indicator
13	34	81.0	10	18	AA08362	Beta-secretase sub
14	34	81.0	10	20	AA033756	Synthetic oligopep
15	34	81.0	10	21	AA069707	Beta-APP alpha-sec
16	34	81.0	10	22	AA010653	Human APP-Sw beta-
17	34	81.0	10	22	AA06898	Human APP-Sw beta-
18	34	81.0	10	22	AA06827	Human amyloid prec
19	34	81.0	10	22	AA007226	Synthetic Asp2 rec
20	34	81.0	10	22	AA007226	Human beta-amyloid
21	34	81.0	10	22	AA047266	Human APP-Sw beta-
22	34	81.0	10	22	AA066575	Peptide 5-5'/SW, fo
23	34	81.0	10	22	AA061337	Synthetic peptide
24	34	81.0	11	22	AA075142	Swedish mutation p
25	34	81.0	11	22	AA075145	APP Swedish varian
26	34	81.0	11	22	AA07469	Asp 1 substrate se
27	34	81.0	16	21	AA060616	Asp2 substrate Swe
28	34	81.0	18	22	AA000609	Human beta-amyloid
29	34	81.0	19	22	AA000611	Beta-amyloid precu
30	34	81.0	20	19	AA082211	Amyloid precursor
31	34	81.0	20	21	AA06714	Fluorogenic protea
32	34	81.0	20	22	AA073229	Beta-APP alpha-sec
33	34	81.0	21	18	AA08361	Protease binding s
34	34	81.0	21	19	AA082186	Beta-secretase sub
35	34	81.0	21	20	AA033755	Synthetic oligopep
36	34	81.0	21	22	AA073203	Protease binding s
37	34	81.0	21	22	AA073204	Asp 1 substrate se
38	34	81.0	21	22	AA047265	Asp2 substrate Swe
39	34	81.0	23	22	AA075148	Beta-secretase sub
40	34	81.0	23	22	AA075147	Synthetic oligopep
41	34	81.0	30	18	AA08360	Substrate for beta
42	34	81.0	30	20	AA033754	Peptide 26-4'/SW, f
43	34	81.0	30	21	AA07895	Mouse amyloid prec
44	34	81.0	30	22	AA047264	
45	34	81.0	32	17	AA004402	

ALIGNMENTS

RESULT 1
AAB07882 standard; peptide; 9 AA.
AAB07882;
14-NOV-2000 (first entry)
A peptide fragment derived from beta-amyloid precursor protein.
Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
inhibitor.
Homo sapiens.
W0200047618-A2.
17-AUG-2000.
10-FEB-2000; 2000WO-US03819.
10-FEB-1999; 99US-0119571.
15-JUN-1999; 99US-0139172.
(ELAN-) ELAN PHARM INC.
Anderson JP, Basti G, Doane MT, Frigon N, John V, Power M;
Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
WPI; 2000-533011/48.
Purified beta-secretase protein used in assays to discover inhibitors
which can be used for the treatment of amyloidogenic diseases e.g. "

PT Alzheimer's disease -
 XX
 PS Disclosure; Page 12; 121pp; English.
 XX
 CC The specification describes a beta-secretase enzyme. The enzyme cleaves
 CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
 CC enzyme is therefore implicated in the production of amyloid plaque
 CC components which accumulate in the brains of individuals afflicted with
 CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
 CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
 CC disease-like pathology to test if they maintain or improve cognitive
 CC ability or reduce the plaque burden. The compounds are used for the
 CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
 CC present sequence represents a peptide derived from beta-amyloid
 CC precursor protein
 CC
 XX Sequence 9 AA:
 SQ
 Query Match 100.0%; Score 42; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SEVNLLAEF 9
 DB 1 SEVNLLAEF 9
 RESULT 2
 AAE00612
 ID AAE00612 standard; peptide; 19 AA.
 XX
 AC AAE00612;
 XX
 DT 02-JUL-2001 (first entry)
 XX
 DE PCBC1 APP beta-secretase cleavage site (D/I) mutant.
 XX
 KW Amyloid precursor protein; APP; caspase; beta-secretase; cleavage site;
 KW cysteine protease; apoptosis; caspase expression cassette; metastasis;
 KW tumour; cathepsin B; urokinase; proliferation; gene therapy;
 KW interdomain linker; Alzheimer's disease; mutant; mutain.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 12 /note= "Wild-type Asp substituted by Ile"
 FT
 XX
 PN WO200129232-A2.
 XX
 PD 26-APR-2001.
 XX
 PF 19-OCT-2000; 2000WO-US28941.
 XX
 PR 20-OCT-1999; 99US-0160559.
 PR 14-AUG-2000; 2000US-0225564.
 XX
 PA (SCIO-) SCIOS INC.
 XX
 PI Cordell B, Li Y;
 XX
 DR WPI; 2001-290920/30.
 XX
 DR Novel fusion polypeptide comprising first and second caspase subunit
 PT separated by cleavage site not associated in nature with caspase
 PT subunit, useful for cloning gene encoding enzymes involved in
 PT proteolytic cleavage -
 XX
 PS Disclosure; Fig 29A; 116pp; English.
 XX
 CC The present amino acid sequence is a PCBC1 plasmid amyloid precursor
 CC protein (APP) beta-secretase cleavage site (AAE00611) (D/I) mutant. APP
 CC sequence containing this mutation is less susceptible to beta-secretase

CC cleavage. APP beta-secretase cleavage site is used to construct an
 CC artificially engineered chimeric cassette comprising human caspase-3 with
 CC interdomain linker replaced by swedish mutant beta-secretase cleavage
 CC site. This modified caspase-3 plays a pivotal role in Alzheimer's
 CC disease. Caspases are a family of cysteine proteases, that participate
 CC in the initiation and execution of apoptosis.
 CC The present invention relates to a method for functional cloning of genes
 CC encoding proteins or enzymes involved in proteolytic cleavage. The
 CC invention is based on the use of caspase expression cassettes comprising
 CC the coding sequence of a proteolytic cleavage site flanked by sequences
 CC encoding two caspase subunits, separated by a cleavage site not associated
 CC and a second caspase subunit, separated by a cleavage site not associated
 CC in nature, is useful for cloning gene encoding enzymes involved in
 CC proteolytic cleavage. An expression cassette containing fusion
 CC polypeptide is used to identify a mutant cell line deficient in an
 CC enzyme of interest and is also useful for diagnosis and suppression of
 CC proliferation or metastases of a tumour cell characterised by
 CC overexpression of a polypeptide (e.g. Cathepsin B or urokinase,
 CC selectively expressed in the tumour cells). DNA encoding fusion
 CC polypeptide is used in gene therapy.
 CC
 XX Sequence 19 AA:
 SQ
 Query Match 95.2%; Score 40; DB 22; Length 19;
 Best Local Similarity 88.9%; Pred. No. 0.12;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SEVNLLAEF 9
 DB 7 SEVNLLAEF 15
 RESULT 3
 AAB07888
 ID AAB07888 standard; peptide; 14 AA.
 XX
 AC AAB07888;
 XX
 DT 14-NOV-2000 (first entry)
 XX
 DE A peptide fragment derived from beta-amyloid precursor protein.
 XX
 KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
 KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
 KW inhibitor.
 KW
 OS Homo sapiens.
 XX
 PN WO200047618-A2.
 XX
 PD 17-AUG-2000.
 XX
 PF 10-FEB-2000; 2000WO-US03819.
 XX
 PR 10-FEB-1999; 99US-0119571.
 PR 15-JUN-1999; 99US-0139172.
 XX
 PA (ELAN-) ELAN PHARM INC.
 XX
 PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
 PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
 XX
 DR WPI; 2000-533011/48.
 XX
 DR Purified beta-secretase protein used in assays to discover inhibitors
 PT which can be used for the treatment of amyloidogenic diseases e.g.
 PT Alzheimer's disease -
 XX
 PS Disclosure; Page 12; 121pp; English.
 XX
 CC The specification describes a beta-secretase enzyme. The enzyme cleaves
 CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
 CC enzyme is therefore implicated in the production of amyloid plaque


```
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC precursor sequence represents a peptide derived from beta-amyloid
CC precursor protein
XX
SQ Sequence 14 AA;

Query Match 92.9%; Score 39; DB 21; Length 14;
Best Local Similarity 88.9%; Pred. No. 0.14;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SEVNLAEF 9
   |||||
Db 6 SEVNLVAEF 14

RESULT 4
AAB07887
ID AAB07887 standard; peptide; 9 AA.
XX
AC AAB07887;
XX
DT 14-NOV-2000 (first entry)
XX
DE A peptide fragment derived from beta-amyloid precursor protein.
XX
KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
XX inhibitor.
XX
OS Homo sapiens.
XX
PN WO200047618-A2.
XX
PD 17-AUG-2000.
XX
PF 10-FEB-2000; 2000WO-US03819.
XX
PR 10-FEB-1999; 99US-0119571.
XX
PR 15-JUN-1999; 99US-0139172.
XX
PA (ELAN-) ELAN PHARM INC.
XX
PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX WPI: 2000-533011/48.
XX
PT Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
PS Disclosure; Page 12; 121pp; English.
XX
CC The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a peptide derived from beta-amyloid
CC precursor protein
XX
SQ Sequence 9 AA;

Query Match 90.5%; Score 38; DB 21; Length 9;
Best Local Similarity 88.9%; Pred. No. 0.14;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SEVNLAEF 9
   |||||
Db 1 SEVNLVAEF 9

RESULT 5
AAB07879
ID AAB07879 standard; peptide; 9 AA.
XX
AC AAB07879;
XX
DT 14-NOV-2000 (first entry)
XX
DE A peptide fragment derived from beta-amyloid precursor protein.
XX
KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
XX inhibitor.
XX
OS Homo sapiens.
XX
PN WO200047618-A2.
XX
PD 17-AUG-2000.
XX
PF 10-FEB-2000; 2000WO-US03819.
XX
PR 10-FEB-1999; 99US-0119571.
XX
PR 15-JUN-1999; 99US-0139172.
XX
PA (ELAN-) ELAN PHARM INC.
XX
PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX WPI: 2000-533011/48.
XX
PT Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
PS Disclosure; Page 12; 121pp; English.
XX
CC The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a peptide derived from beta-amyloid
CC precursor protein
XX
SQ Sequence 9 AA;

Query Match 88.1%; Score 37; DB 21; Length 9;
Best Local Similarity 88.9%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLAEF 9
   |||||
Db 1 SEVNLVAEF 9

RESULT 6
AAB07883
ID AAB07883 standard; peptide; 9 AA.
XX
```

AC	AAB07883.
XX	14-NOV-2000 (first entry)
DT	
XX	A peptide fragment derived from beta-amyloid precursor protein.
DE	
KW	Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KM	amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW	inhibitor.
XX	
OS	Homo sapiens.
XX	
PN	WO200047618-A2.
XX	
PD	17-AUG-2000.
XX	
PF	10-FEB-2000; 2000WO-US03819.
XX	
PR	10-FEB-1999; 99US-0119571.
PR	15-JUN-1999; 99US-0139172.
XX	
PA	(ELAN-) ELAN PHARM INC.
XX	
PI	Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI	Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX	
DR	WPI; 2000-533011/48.
XX	
PT	Purified beta-secretase protein used in assays to discover inhibitors
PT	which can be used for the treatment of amyloidogenic diseases e.g.
PT	Alzheimer's disease -
XX	
XX	Disclosure; Page 12; 121pp; English.
XX	
CC	The specification describes a beta-secretase enzyme. The enzyme cleaves
CC	beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC	enzyme is therefore implicated in the production of amyloid plaque
CC	components which accumulate in the brains of individuals afflicted with
CC	Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC	a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC	disease-like pathology to test if they maintain or improve cognitive
CC	ability or reduce the plaque burden. The compounds are used for the
CC	treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC	present sequence represents a peptide derived from beta-amyloid
CC	precursor protein
XX	
SQ	Sequence 9 AA:
XX	
Query Match	85.7%; Score 36; DB 21; Length 9;
Best Local Similarity	88.9%; Pred. NO. 6.4e+05;
Matches 8; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1 SEVNLAEF 9
DB	11111111 1 SEVKLLAEF 9
RESULT 7	
ID ABB64111	
ABB64111 standard; Protein: 649 AA.	
AC ABB64111;	
XX	
DT 26-MAR-2002 (first entry)	
XX	
DE Drosophila melanogaster polypeptide SEQ ID NO 19125.	
XX	
KW Drosophila; developmental biology; cell signalling; insecticide;	
KM pharmaceutical.	
XX	
OS Drosophila melanogaster.	
XX	
PN WO200171042-A2.	

XX	27-SEP-2001.
PD	
XX	
PE	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
XX	
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
DR	WPI: 2001-656860/75.
N-PSDB:	ABU08214.
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
PT	
XX	
XX	Disclosure: SEQ ID NO 19125; 21pp + Sequence Listing; English.
CC	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABB57737-ABB72072).
CC	((ABB57737-ABB72072)).
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
CC	
CC	
XX	
SQ	Sequence 649 AA:
OY	1 SEVNILAE 8 : Query Match 83.3%; Score 35; DB 22; Length 649; Best Local Similarity 87.5%; Pred. No. 60; Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db	131 SEINLAE 138
RESULT 8	
AAM82081	
ID	AAM82081 standard; peptide: 9 AA.
XX	
AC	AAM82081;
XX	
DIT	18-FEB-1999 (first entry)
DE	Fluorogenic protease indicator protease binding peptide #59.
XX	
KW	Protease activity; fluorophore; detection; fluorogenic; cellular uptake; conformation change.
XX	
OS	Synthetic.
PN	WO9837226-A1.
XX	
PD	27-AUG-1998.
XX	
PF	20-FEB-1998; 98WO-US03000.
XX	
PR	20-FEB-1997; 97US-0802981.
XX	
PA	(ONCO-) ONCOTIMMUNIN INC.
XX	
PI	Komoriya A, Packard BS;
XX	
DR	WPI, 1998-467579/40.
XX	

PT New fluorogenic compositions - containing 2 fluorophores separated
PT by a peptide comprising a protease binding site, used for detecting
XX protease activity in samples.
XX
XX Claim 4; Page 77; 90pp; English.
XX
CC AAM82023-W82240 are peptides used in the construction of a fluorogenic
CC composition which is used for the detection of protease activity in
CC biological samples. The products can be used for the detection of
CC conformation changes in nucleic acids, oligosaccharides,
CC polysaccharides, proteins or polymers. In addition, attachment of a
CC glycoproteins, steroids or lipids, phospholipids, glycolipids,
CC hydrophobic group to a molecule can be used to enhance uptake by cells.
CC The composition is composed of P - peptide comprising a protease binding
CC site for the protease, F1, F2 peptides - fluorophores where F1 is
CC attached to the amino terminal amino acid and F2 is attached to the
CC carboxyl terminal amino acid and S1, S2 peptides - when present, are
CC peptide spacers where S1, when present, is attached to the amino terminal
CC acid, and S2, when present, is attached to the carboxyl terminal amino
CC acid.
XX
SQ Sequence 9 AA:
XX
Query Match 81.0%; Score 34; DB 19; Length 9;
Best Local Similarity 88.9%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SEVNLLAEF 9
DB 1 SEVNLLAEF 9
XX
RESULT 9
AAB07874
ID AAB07874 standard; peptide; 9 AA.
XX
AC AAB07874;
XX
DT 14-NOV-2000 (first entry)
XX
DE A peptide fragment derived from beta-amyloid precursor protein.
XX
KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
XX inhibitor.
XX
OS Homo sapiens.
XX
PN WO200047618-A2.
XX
PD 17-AUG-2000.
XX
PF 10-FEB-2000; 2000WO-US03819.
XX
PR 10-FEB-1999; 99US-0119571.
PR 15-JUN-1999; 99US-0139172.
XX
PA (ELAN-) ELAN PHARM INC.
XX
PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Slnha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX
DR WPI; 2000-533011/48.
XX
PT Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
XX Alzheimer's disease -
XX
PS Disclosure; Page 12; 121pp; English.
XX
CC The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a peptide derived from beta-amyloid
CC precursor protein
XX
SQ Sequence 9 AA:
XX
Query Match 81.0%; Score 34; DB 19; Length 9;
Best Local Similarity 88.9%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a peptide derived from beta-amyloid
CC precursor protein
XX
SQ Sequence 9 AA:
XX
Query Match 81.0%; Score 34; DB 21; Length 9;
Best Local Similarity 88.9%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SEVNLLAEF 9
DB 1 SEVNLLAEF 9
XX
RESULT 10
AAB07881
ID AAB07881 standard; peptide; 9 AA.
XX
AC AAB07881;
XX
DT 14-NOV-2000 (first entry)
XX
DE A peptide fragment derived from beta-amyloid precursor protein.
XX
KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
XX inhibitor.
XX
OS Homo sapiens.
XX
PN WO200047618-A2.
XX
PD 17-AUG-2000.
XX
PF 10-FEB-2000; 2000WO-US03819.
XX
PR 10-FEB-1999; 99US-0119571.
PR 15-JUN-1999; 99US-0139172.
XX
PA (ELAN-) ELAN PHARM INC.
XX
PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Slnha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX
DR WPI; 2000-533011/48.
XX
PT Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
XX Alzheimer's disease -
XX
PS Disclosure; Page 12; 121pp; English.
XX
CC The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a peptide derived from beta-amyloid
CC precursor protein
XX
SQ Sequence 9 AA:
XX
Query Match 81.0%; Score 34; DB 21; Length 9;

Best Local Similarity 77.8%; Pred. No. 6.4e+05;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
| | | : | | |
Db 1 SEVNLAEF 9

RESULT 11

AAB07894 standard; Peptide: 9 AA.

AAB07894;

14-NOV-2000 (first entry)

Substrate for beta-secretase enzyme.

Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
inhibitor.

Synthetic.

Key Location/Qualifiers
Cleavage-site 5..6

WO200047618-A2.

17-AUG-2000.

10-FEB-2000; 2000WO-US03819.

10-FEB-1999; 99US-0119571.
15-JUN-1999; 99US-0139172.

(ELAN-) ELAN PHARM INC.

Anderson JP, Basl G, Doane MT, Frigon N, John V, Power M;
Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;

WPI: 2000-533011/48.

Purified beta-secretase protein used in assays to discover inhibitors
which can be used for the treatment of amyloidogenic diseases e.g.

Alzheimer's disease -

Example 4; Page 71; 121pp; English.

The specification describes a beta-secretase enzyme. The enzyme cleaves
beta-amyloid precursor protein to produce beta-amyloid peptide. This
enzyme is therefore implicated in the production of amyloid plaque
components which accumulate in the brains of individuals afflicted with
Alzheimer's disease. Inhibitors of beta-secretase are administered to
a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
disease-like pathology to test if they maintain or improve cognitive
ability or reduce the plaque burden. The compounds are used for the
treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
present sequence represents a peptide substrate used to test the
activity of beta-secretase enzyme.

Sequence 9 AA;

Query Match 81.0%; Score 34; DB 21; Length 9;

Best Local Similarity 88.9%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
| | | | | | | | |
OS

Db 1 SEVNLAEF 9

RESULT 12

AAG73297
ID AAG73297 standard; Peptide: 9 AA.

AAG73297;

14-AUG-2001 (first entry)

Protease indicator compound peptide #26.

Protease detection; peptide cleavage; enzyme activity; fluorogenic;
viral infection; cancer metastasis; emphysema; arthritis;
thrombosis; haemophilia.

Synthetic.

WO200118238-A1.

11-SEP-2000; 2000WO-US24882.

10-SEP-1999; 99US-0394019.

(ONCO-) ONCOIMMUNIN INC.

Komoriya A, Packard BS;

WPI: 2001-389573/41.

New fluorogenic compositions whose fluorescence level increases in the
presence of active proteases, useful for detecting and localizing
protease activity in biological samples, particularly in frozen tissue
samples -

Claim 1; Page 70; 86pp; English.

The present invention describes fluorogenic compositions which can be
used for the detection of protease activity. This can be useful as an
indicator of viral infection, cancer metastasis, haemophilia, emphysema,
thrombosis and arthritis. The fluorogenic compositions comprise a
peptide, a peptide spacer and a donor and an acceptor fluorophore. The
peptide is cleaved by a protease and the fluorophores can then be
detected. The present sequence is one of the peptides described in the
exemplification of the invention.

Sequence 9 AA;

Query Match 81.0%; Score 34; DB 22; Length 9;

Best Local Similarity 88.9%; Pred. No. 6.4e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLAEF 9
| | | | | | | | |
OS

Db 1 SEVNLAEF 9

RESULT 13

AAW08362
ID AAW08362 standard; peptide: 10 AA.

AAW08362;

05-SEP-1997 (first entry)

Beta-secretase substrate #3.

Beta-cleavage site; beta amyloid precursor protein; APP; beta-secretase;
alpha-secretase; proteolytic cleavage; inhibitor; Alzheimer's disease.

Synthetic.

Key Location/Qualifiers

Misc-difference 1

FT /note- "acetylated"

XX
PN W09640885-A2.
XX
PD 19-DEC-1996.
XX
PE 07-JUN-1996; 96WO-US09985.
XX
PR 07-JUN-1995; 95US-0485152.
PR 07-JUN-1995; 95US-0480498.
XX
PA (ATHE-) ATHENA NEUROSCIENCES INC.
XX
PI Anderson JP, Chrysler SMS, Jacobson-croak KL, Keim PS;
PI Mcconlogue LC, Sinha S, Tan H;
DR WPI; 1997-052304/05.
XX
PT Beta-secretase which specifically cleaves beta-amyloid precursor
PT protein - useful to screen for inhibitors useful in treatment of
PT Alzheimer's disease
XX
PS Disclosure; Page 45; 92pp; English.
XX
CC AAW08359-W08362 represent substrates for the enzyme of the invention.
CC The enzyme of the invention is beta-secretase, and specifically cleaves
CC beta-amyloid precursor protein (beta-APP). Normal processing of beta-APP
CC is thought to occur via cleavage between residues 16 and 17 of the
CC beta-amyloid peptide region by an alpha-secretase. Pathogenic processing
CC is thought to occur by beta-secretase cleavage of beta-APP.
CC Beta-secretase activity can be detected and measured using a method of
CC the invention, which detects at least one of the beta-secretase cleavage
CC products formed on cleavage. The method can be used to determine whether
CC a test substance inhibits proteolytic cleavage, by beta-secretase, of
CC beta-APP. Compounds effective to at least partially inhibit
CC beta-secretase activity can be used to inhibit cleavage of beta-APP in
CC cells or mammalian hosts. Isolation and purification of beta-secretase
CC will permit chemical modelling of a critical event in the pathology of
CC Alzheimer's disease.
XX
SQ Sequence 10 AA:

Query Match 81.0%; Score 34; DB 18; Length 10;
Best Local Similarity 88.9%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAEF 9
DB 1 SEVNLLAEF 9

RESULT 14
AAV33756
ID AAV33756 standard; Protein; 10 AA.
XX
AC AAV33756;
XX
DT 09-NOV-1999 (first entry)
XX
DE Synthetic oligopeptide 5-5'SW.
XX
KW Beta-secretase; beta-amyloid protein precursor; APP; Down's syndrome;
KW Alzheimer's disease; measure activity; cleavage site.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Modified-site 1 /note- "N-terminal Ser is acetylated"
XX
XX US5942400-A.
XX
XX 24-AUG-1999.

XX
PE 07-JUN-1996; 96US-0659984.
XX
PR 07-JUN-1996; 96US-0659984.
PR 07-JUN-1995; 95US-0480498.
PR 07-JUN-1995; 95US-0485152.
XX
PA (ELAN-) ELAN PHARM INC.
XX
PI Anderson JP, Jacobson-Croak KL, Sinha S;
PI WPI; 1999-517417/43.
DR
XX
PT A method for detecting human beta-secretase cleavage of polypeptides
PT useful for identifying beta-secretase inhibitors
XX
PS Examples; Column 30; 43pp; English.
XX
CC Sequences AAV33752-Y33756 are synthetic oligopeptides used for measuring
CC the activity of beta-secretase (AAV33741). Beta-secretase is capable of
CC cleaving beta-amyloid protein precursor (APP) (AAV33742). These
CC synthetic peptides contain the cleavage site of APP. Beta-secretase and
CC APP are used in a method for detecting human beta-secretase cleavage of
CC polypeptides and for identifying beta-secretase inhibitors. Inhibition
CC of beta-secretase activity would be useful for chemical modelling of a
CC critical event in the pathology of Alzheimer's disease. Inhibitors of
CC beta-secretase would be useful for the prevention and treatment of
CC Alzheimer's disease and Down's Syndrome.
XX
SQ Sequence 10 AA:

Query Match 81.0%; Score 34; DB 20; Length 10;
Best Local Similarity 88.9%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SEVNLLAEF 9
DB 1 SEVNLLAEF 9

RESULT 15
AAV69707
ID AAV69707 standard; peptide; 10 AA.
XX
AC AAV69707;
XX
DT 11-APR-2000 (first entry)
XX
DE Beta-APP alpha-secretase substrate [NMD]-APP(-5,+5).
XX
KW Neotropic; neuroprotective; beta-amyloid precursor protein; metabolism;
KW cleavage site; beta-secretase; neurodegenerative disease;
KW Alzheimer's disease.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN W09964587-A1.
XX
PD 16-DEC-1999.
XX
PE 04-JUN-1999; 99WO-FR01326.
XX
PR 05-JUN-1998; 98FR-0007068.
PR 31-MAR-1999; 99US-0122599.
XX
PA (RHON) RHONE-POULENC RORER SA.
PA (UYPA-) UNIV CURIE PARIS VI P & M.
XX
PI Rholam M, Munoz-Gimenez N, Moutaouakil M, Cohen P, Bertrand P;
PI WPI; 2000-097537/08.
DR
XX

PT Polypeptide with beta-secretase activity, specific for wild-type
PT amyloid precursor protein, useful in treating Alzheimer's disease

PS Example 3; Page 24; 44pp; French.

XX Peptides AAY69702-Y69718 represent synthetic peptide substrates for a
CC novel polypeptide with beta-secretase activity that can cleave
CC specifically the natural beta-amyloid precursor protein (bAPP). Normal
CC cleavage of the protein occurs between amino acids Met596-Asp597 and
CC Val636-Ile637 (positions 4-5 and 44-45 of AAY69701). The novel
CC polypeptide is used to identify agents that interact specifically with
CC it. These agents regulate metabolism of APP, particularly they slow down
CC or reduce production of beta-amyloid, so can be used to treat
CC neurodegenerative diseases, particularly Alzheimer's disease.

XX Sequence 10 AA;

Query Match 81.0%; Score 34; DB 21; Length 10;

Best Local Similarity 88.9%; Pred. No. 1;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SEVNLDAEF 9

DB 1 SEVNLDAEF 9

Search completed: October 30, 2002, 12:27:09
Job time : 3.60442 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 30, 2002, 12:24:20 ; Search time 51.4447 Seconds

(without alignments)
171.400 Million cell updates/sec

Title: US-09-724-571-75

Perfect score: 1807

Sequence: 1 MVDLNRKSGSGGYVETVG.....VFEDRRKRIGFVNSACHVH 361

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

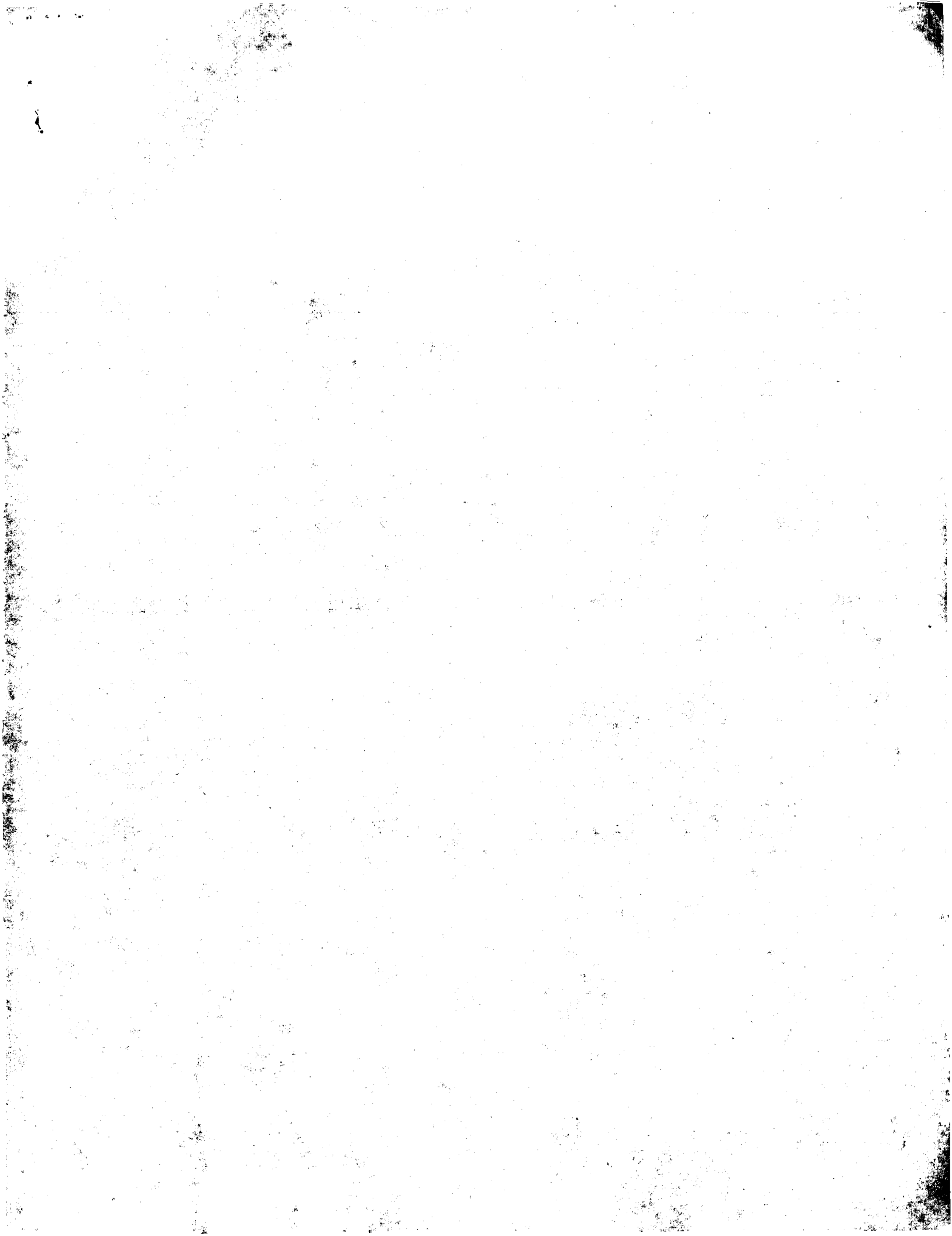
SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1901	99.7	501	4	US-09-009-191-2 Sequence 2, Appl1
2	1890	99.1	774	4	US-09-009-191-4 Sequence 4, Appl1
3	1106	58.0	518	3	US-08-999-723-2 Sequence 2, Appl1
4	1106	58.0	518	4	US-09-434-427-2 Sequence 2, Appl1
5	1088	57.1	514	4	US-09-912-484-2 Sequence 2, Appl1
6	1088	57.1	514	4	US-08-208-007A-13 Sequence 13, Appl1
7	298.5	15.7	396	4	US-08-032-523-9 Sequence 9, Appl1
8	298.5	15.7	396	4	US-08-208-007A-12 Sequence 12, Appl1
9	295.5	15.5	412	1	US-08-974-691-4 Patent No. 5217891
10	295.5	15.5	412	4	US-08-974-691-4 Patent No. 5217891
11	279.5	14.7	458	6	Sequence 6, Appl1
12	279.5	14.3	409	1	US-09-640-305-6 Sequence 6, Appl1
13	273.5	14.3	409	1	US-08-360-673-6 Sequence 8, Appl1
14	273.5	14.3	427	2	US-08-846-021A-8 Sequence 2, Appl1
15	270	14.2	410	1	US-08-088-633-2 Sequence 2, Appl1
16	270	14.2	410	1	US-08-088-633-2 Sequence 2, Appl1
17	270	14.2	410	1	US-08-245-756-2 Sequence 2, Appl1
18	270	14.2	410	1	US-08-441-751-2 Sequence 2, Appl1
19	270	14.2	410	1	US-08-441-751-2 Sequence 2, Appl1
20	249	13.1	349	5	PCR-US92-02521-2 Sequence 2, Appl1
21	227	11.9	398	1	US-09-032-523-3 Sequence 2, Appl1
22	227	11.9	398	1	US-08-328-314-2 Sequence 2, Appl1
23	212	11.1	419	4	US-08-974-691-3 Sequence 3, Appl1
24	195.5	10.3	397	3	US-09-079-415-2 Sequence 2, Appl1
25	195.5	10.3	427	1	US-07-958-222A-2 Sequence 2, Appl1
26	194.5	10.2	430	1	US-08-535-237-2 Sequence 2, Appl1
27	193	10.1	330	3	US-08-113-735-1 Sequence 1, Appl1

28	193	10.1	419	3	US-08-115-753-2 Sequence 2, Appl1
29	193	10.1	419	3	US-08-115-753-3 Sequence 33, Appl1
30	184	9.6	445	4	US-08-974-691-6 Sequence 6, Appl1
31	184	9.6	451	4	US-08-974-691-2 Sequence 2, Appl1
32	180	9.4	420	4	US-08-008-271A-4 Sequence 4, Appl1
33	180	9.4	420	4	US-08-974-691-8 Sequence 8, Appl1
34	177.5	9.3	395	2	US-08-723-938-3 Sequence 3, Appl1
35	177.5	9.3	395	2	US-08-080-538-3 Sequence 3, Appl1
36	149	7.8	437	4	US-09-353-332-2 Sequence 2, Appl1
37	129.5	6.8	140	3	US-09-211-631-13 Sequence 13, Appl1
38	129.5	6.8	140	4	US-09-265-628-13 Sequence 13, Appl1
39	129.5	6.8	140	4	US-09-001-141-11 Sequence 11, Appl1
40	129.5	6.8	140	4	US-09-532-803-6 Sequence 6, Appl1
41	129.5	6.8	140	4	US-09-653-403-14 Sequence 14, Appl1
42	97	5.1	1030	4	US-09-091-117-2 Sequence 2, Appl1
43	95.5	5.0	280	4	US-09-160-246-14 Sequence 14, Appl1
44	85	4.5	377	2	US-08-853-659A-41 Sequence 41, Appl1
45	84.5	4.4	1097	2	US-08-680-326-39 Sequence 39, Appl1

ALIGNMENTS

RESULT 1
US-09-009-191-2
Sequence 2, Application US/09009191
Patent No. 6319689
GENERAL INFORMATION:
APPLICANT: POWELL, DAVID
APPLICANT: CHAPMAN, CONRAD
APPLICANT: MORPHY, KAY
APPLICANT: SMITH, TRUDI
TITLE OF INVENTION: ASP2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,191
FILING DATE: 20-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 9701684.4
FILING DATE: 128-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70368
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-191-2
Query Match 99.7%; Score 1901; DB 4; Length 501;
Best Local Similarity 99.7%; Pred. No. 1.4e-187;
Matches 360; Conservative 0; Mismatches 1; Indels 0; Gaps 0;




```

QY 1 MYDNLGKSGGCGIYVEMTVGSPPTLNILVDTGSSNFAVGAAPHFLHRYRQRLSSTYR 60
Db 63 MYDNLGKSGGCGIYVEMTVGSPPTLNILVDTGSSNFAVGAAPHFLHRYRQRLSSTYR 122
QY 61 DLKRGVYPTGKMEGELGDLVSIPIHGPVNTVANAIAATESDKFFINSNMEGILGL 120
Db 123 DLKRGVYPTGKMEGELGDLVSIPIHGPVNTVANAIAATESDKFFINSNMEGILGL 182
QY 121 AYAEIARPPDLSLEPFDSLVKQTHVPLNLSIQLGAGFPLNQSIVLASVSGSMITIGIDH 180
Db 183 AYAEIARPPDLSLEPFDSLVKQTHVPLNLSIQLGAGFPLNQSIVLASVSGSMITIGIDH 242
QY 181 SLYTSLSMTYPIRRRWYEVILVREINIGODLKMCKEYNDKSIYDSGTTNLRPKV 240
Db 243 SLYTSLSMTYPIRRRWYEVILVREINIGODLKMCKEYNDKSIYDSGTTNLRPKV 302
QY 241 EAAVKSIRAASTKTEKPDGFVLGEOVLVCMQAGTTPMNIPIVLSIYLMGEVNTOSFRITIL 300
Db 303 EAAVKSIRAASTKTEKPDGFVLGEOVLVCMQAGTTPMNIPIVLSIYLMGEVNTOSFRITIL 362
QY 301 POQYLRPEVATSDODCKYKFAISQSTGYVMGAVIMEGFYVDFDRAKRKIGFAVSACHV 360
Db 363 POQYLRPEVATSDODCKYKFAISQSTGYVMGAVIMEGFYVDFDRAKRKIGFAVSACHV 422
QY 361 H 361
Db 423 H 423

```

RESULT 2
US-09-009-191-4
Sequence 4, Application US/09009191
Patent No. 6319689

GENERAL INFORMATION:
APPLICANT: POWELL, DAVID
APPLICANT: CHAPMAN, CONRAD
APPLICANT: MORPHY, KAY
APPLICANT: SMITH, TRUDI
TITLE OF INVENTION: ASP2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,191
FILING DATE: 20-JAN-1998
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: UK 9701684.4
FILING DATE: 28-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70368
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 amino acids
TYPE: amino acid
RANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-191-4

Query Match 99.1%; Score 1890; DB 4; Length 774;
Best Local Similarity 99.2%; Pred. No. 3,76-186;
Matches 358; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 MYDNLGKSGGCGIYVEMTVGSPPTLNILVDTGSSNFAVGAAPHFLHRYRQRLSSTYR 60
Db 6 MYDNLGKSGGCGIYVEMTVGSPPTLNILVDTGSSNFAVGAAPHFLHRYRQRLSSTYR 65
QY 61 DLKRGVYPTGKMEGELGDLVSIPIHGPVNTVANAIAATESDKFFINSNMEGILGL 120
Db 66 DLKRGVYPTGKMEGELGDLVSIPIHGPVNTVANAIAATESDKFFINSNMEGILGL 125
QY 121 AYAEIARPPDLSLEPFDSLVKQTHVPLNLSIQLGAGFPLNQSIVLASVSGSMITIGIDH 180
Db 126 AYAEIARPPDLSLEPFDSLVKQTHVPLNLSIQLGAGFPLNQSIVLASVSGSMITIGIDH 185
QY 181 SLYTSLSMTYPIRRRWYEVILVREINIGODLKMCKEYNDKSIYDSGTTNLRPKV 240
Db 186 SLYTSLSMTYPIRRRWYEVILVREINIGODLKMCKEYNDKSIYDSGTTNLRPKV 245
QY 241 EAAVKSIRAASTKTEKPDGFVLGEOVLVCMQAGTTPMNIPIVLSIYLMGEVNTOSFRITIL 300
Db 246 EAAVKSIRAASTKTEKPDGFVLGEOVLVCMQAGTTPMNIPIVLSIYLMGEVNTOSFRITIL 305
QY 301 POQYLRPEVATSDODCKYKFAISQSTGYVMGAVIMEGFYVDFDRAKRKIGFAVSACHV 360
Db 306 POQYLRPEVATSDODCKYKFAISQSTGYVMGAVIMEGFYVDFDRAKRKIGFAVSACHV 365
QY 361 H 361
Db 366 H 366

```

RESULT 3
US-08-999-723-2
Sequence 2, Application US/08999723A
Patent No. 6025180

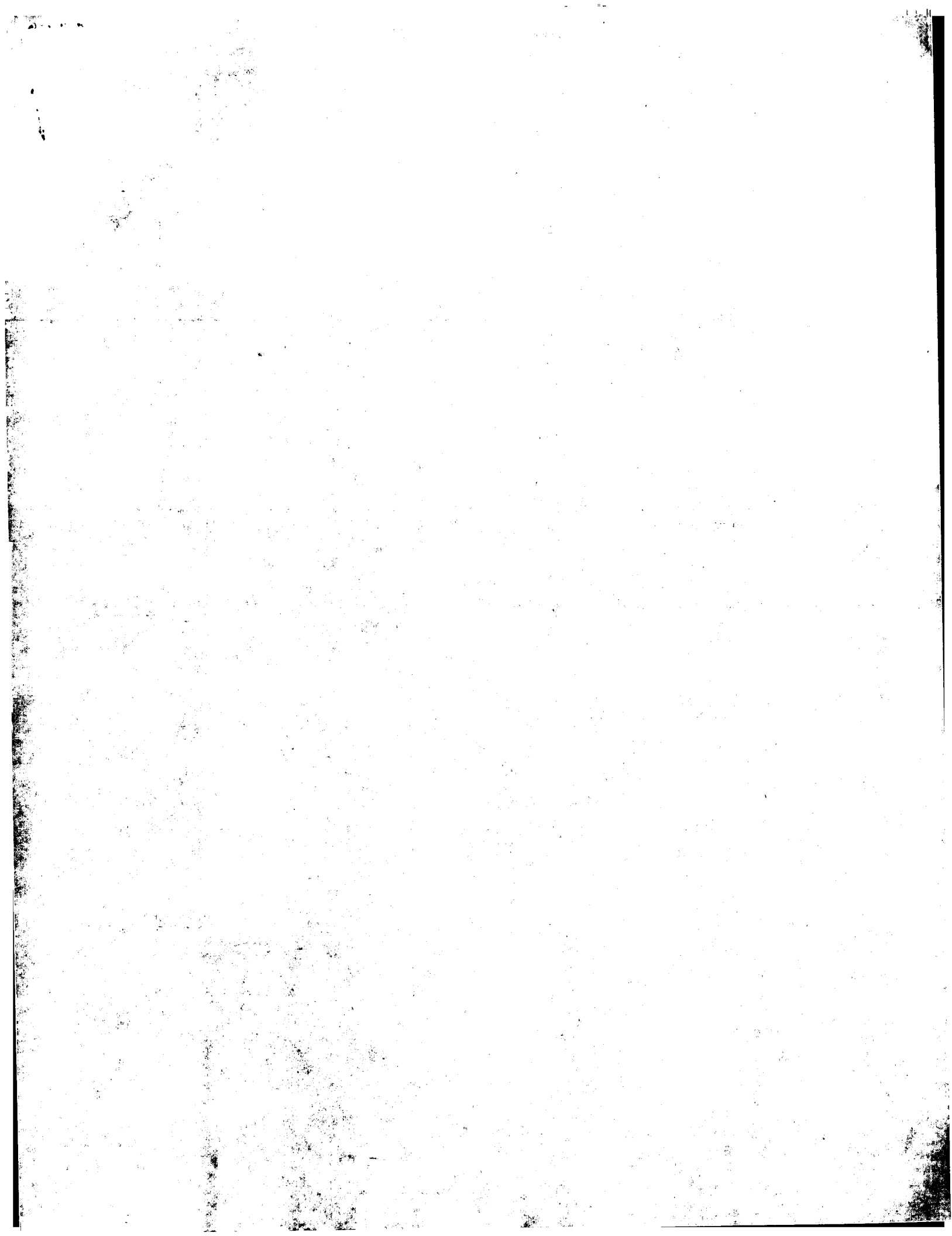
GENERAL INFORMATION:
APPLICANT: Powell, David J.
APPLICANT: Soutchan, Christopher
APPLICANT: Chapman, Conrad G.
APPLICANT: Evans, Joane R.
TITLE OF INVENTION: ASP1
FILE REFERENCE: GH70262
CURRENT APPLICATION NUMBER: US/08/999,723A
CURRENT FILING DATE: 1997-10-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 518
TYPE: PRT
ORGANISM: Homo sapiens
US-08-999-723-2

Query Match 58.0%; Score 1106; DB 3; Length 518;
Best Local Similarity 56.4%; Pred. No. 1,5e-105;
Matches 202; Conservative 60; Mismatches 92; Indels 4; Gaps 2;

```

QY 1 MYDNLGKSGGCGIYVEMTVGSPPTLNILVDTGSSNFAVGAAPHFLHRYRQRLSSTYR 60
Db 80 MYDNLGKSGGCGIYVEMTVGSPPTLNILVDTGSSNFAVGAAPHFLHRYRQRLSSTYR 139
QY 61 DLKRGVYPTGKMEGELGDLVSIPIHGPVNTVANAIAATESDKFFINSNMEGILGL 120
Db 140 SKGFDTVAKYITGSGWTGPEGLVITIPKFNPSLVNATATFESBNFPLPIKRWNGITGL 199
QY 121 AYAEIARPPDLSLEPFDSLVKQTHVPLNLSIQLGAGFPLNQSIVLASVSGSMITIGIDH 180
Db 200 AYATLAKPSSSLETFEFDLSIVQANIPNVFSMQMCGAGLPAVGS---GINGSLVLDGIEP 256

```



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2002, 09:16:14 ; Search time 56.55 Seconds

(Without alignments)
799.417 Million cell updates/sec

Title: US-09-724-571-58

Perfect score: 2156
Sequence: 1 ETDEPPEPGRGRGFEVMD.....GPFVTLDMEDCGYNIPQDE 407

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
A.Geneseq_032802:*
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	2156	100.0	415	21	AA807899
2	2156	100.0	453	21	AA884838
3	2156	100.0	453	22	AAE10642
4	2156	100.0	453	22	AAE06872
5	2156	100.0	453	22	AAU06616
6	2156	100.0	453	22	AAU07215
7	2156	100.0	453	22	AAE02594
8	2156	100.0	456	21	AA807897
9	2156	100.0	459	21	AA884839
10	2156	100.0	459	22	AAE10643
11	2156	100.0	459	22	AAE06873

12	2156	100.0	459	22	AAU06617	Human-pro-Asp 2(a)
13	2156	100.0	459	22	AAU07216	Human aspartyl pro
14	2156	100.0	459	22	AAE02595	Human-asp-2(a) del
15	2156	100.0	460	21	AA807898	Amino acid sequenc
16	2156	100.0	488	22	AA865572	Human memapsin 2.
17	2156	100.0	488	22	AA861334	Memapsin 2 protein
18	2156	100.0	501	21	AA94767	Human beta-secreta
19	2156	100.0	501	21	AA807896	Amino acid sequenc
20	2156	100.0	503	22	AA866573	Human pro-memapsin
21	2156	100.0	503	22	AA861335	T7 promoter and ve
22	2151	99.8	433	21	AA884833	Human-pro-asp-2(a)
23	2151	99.8	433	22	AAE10640	Human-pro-asp 2(a)
24	2151	99.8	433	22	AAE06870	Human-pro-asp2(a)
25	2151	99.8	433	22	AAU06614	Human-pro-asp 2(a)
26	2151	99.8	433	22	AAU07213	T7-human aspartyl
27	2151	99.8	433	22	AAE02592	Human-pro-asp-2(a)
28	2151	99.8	446	21	AA884831	T7-caspase-human-p
29	2151	99.8	446	22	AAE10638	T7-human-pro-asp 2
30	2151	99.8	446	22	AAE06868	Human T7-Human-pro
31	2151	99.8	446	22	AAU06612	T7-human aspartyl
32	2151	99.8	446	22	AAU07211	T7-caspase-human-p
33	2151	99.8	446	22	AAE02590	T7-caspase-human-p
34	2151	99.8	459	21	AA884832	T7-caspase-Human-p
35	2151	99.8	459	22	AAE10639	Human T7-Caspase-H
36	2151	99.8	459	22	AAE06869	T7-caspase-human-p
37	2151	99.8	459	22	AAU06613	T7-caspase-human-p
38	2151	99.8	459	22	AAU07212	T7-caspase-human-p
39	2151	99.8	459	22	AAE02591	T7-caspase-human-p
40	2151	99.8	501	21	AA884825	Human aspartyl pro
41	2151	99.8	501	22	AAE10629	Human aspartyl pro
42	2151	99.8	501	22	AAE06859	Human aspartyl pro
43	2151	99.8	501	22	AAU06603	Human aspartyl pro
44	2151	99.8	501	22	AAU07202	Human aspartyl pro
45	2151	99.8	501	22	AAE02581	Human aspartyl pro

ALIGNMENTS

RESULT 1	AA807899	standard; Protein; 415 AA.
XX	AA807899;	
XX	14-NOV-2000	(first entry)
XX	AC	Amino acid sequence of a human beta-secretase enzyme fragment.
XX	DT	Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
XX	XX	amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW	KW	inhibitor.
XX	XX	
OS	OS	Homo sapiens.
XX	XX	
PN	PN	W0200047618-A2.
XX	PD	17-AUG-2000.
XX	PF	10-FEB-2000; 2000MO-US03819.
XX	PR	10-FEB-1999; 9905-0119571.
XX	PR	15-JUN-1999; 9905-0139172.
PA	PA	(ELAN-) ELAN PHARM INC.
XX	PI	Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
XX	PI	Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX	DR	WPI: 2000-533011/48.
XX	PT	Purified beta-secretase protein used in assays to discover inhibitors
PT	PT	which can be used for the treatment of amyloidogenic diseases e.g.

XX	10-DEC-2001	(first entry)
XX		
DE	Human-Asp 2(a)	protein lacking transmembrane domain.
XX		
KM	Alzheimer's disease 2a: Asp2a;	amyloid precursor protein; APP;
KM	Alzheimer's disease; AD;	dementia; neurofibrillary tangle; gliosis;
XX	amyloid plaque; neuronal loss;	proteolytic; neurotrophic; neuroprotective.
XX		
OS	Homo sapiens.	
OS	Synthetic.	
FT	Key	Location/Qualifiers
FT	Misc-difference 214	
XX		/note= "Encoded by CAC"
XX		
PN	GB2357767-A.	
XX		
PD	04-JUL-2001.	
XX		
PF	22-SEP-2000;	2000GB-0023315.
XX		
PR	23-SEP-1999;	99US-0155493.
PR	23-SEP-1999;	99US-0404133.
PR	23-SEP-1999;	99MO-US20881.
PR	13-OCT-1999;	99US-0416901.
PR	06-DEC-1999;	99US-0169232.
XX		
PA	(PRAA)	PHARMACIA & UPJOHN CO.
XX		
PI	Blenkowski MJ,	Gurney M;
XX		
DR	WPI: 2001-444208/48.	
XX		
XX	N-PSDB: AAD17878.	
PT	Polypeptide comprising fragments of human aspartyl protease with	
PT	amyloid precursor protein processing activity and alpha-secretase	
PT	activity, for identifying modulators useful in treating Alzheimer's	
XX	disease -	
XX		
PS	Example 10; Fig 11; 187bp; English.	
XX		
CC	The patent discloses human aspartyl protease 1 (hu-Asp1) or modified	
CC	Asp1 proteins which lack transmembrane domain or amino terminal	
CC	domain or cytoplasmic domain and retains alpha-secretase activity	
CC	and amyloid protein precursor (APP) processing activity. The proteins	
CC	of the invention are useful for assaying hu-Asp1 alpha-secretase	
CC	activity, which in turn is useful for assaying hu-Asp1 alpha-secretase	
CC	hu-Asp1 alpha-secretase activity, where modulators that increase	
CC	hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's	
CC	disease (AD) which causes progressive dementia with consequent	
CC	formation of amyloid plaques, neurofibrillary tangles, gliosis and	
CC	neuronal loss. Hu-Asp1 protease substrate is useful for assaying	
CC	hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with	
CC	the substrate under acidic conditions and determining the level of	
CC	hu-Asp1 proteolytic activity. The present sequence is human Asp 2(a)	
CC	protein lacking a transmembrane (TM) domain. This sequence is generated	
CC	by the deletion of the C-terminal TM domain of human Asp 2(a) protein.	
XX		
XX	Sequence 453 AA;	

QY	121	SKKEFLNCSNMWCGIILGLAYAEIARPDSDLEPPEDSLVKQIYHVNLFSLQICGAGFPLNOS	180
Db	166	sdktfingsnwegllglalayaetarpdsllepfdslvxqthvpmllfsjqlcgagfplngs	223
QY	181	EVLASVGSGLIIGSDHSLYTGSLWYTPYIRREWWYEVYIVVEINSGDLKMDCKEYWDK	240
Db	226	evlaevgsmliggladhslygslwytprlrrewyevylivveingdqlkmdckeywdk	285
QY	241	SLVDSGTTNLRLEPKKVEFAAVKSLKAASSTKPEPDGFWLGEOLYCWQAGTTPWNIFFVIS	300
Db	286	slvdsqctnllrpkpkvefaavkslkaasstekfpgdglwqgagfctppmllffvis	345
QY	301	LYLMEVYNOSFRITILPQOYLREVEDATSDODDCKYFAIISQSGTGVKNAVIMEGYVYV	360
Db	346	lylmgvevngsfritllppqglrpvedadausqddcykflaigsqstglvnmavimegyfyyv	405
QY	361	FDRARRKIGFAVSAACHVDEFRIAAVEGPFVYTLDMECGNIIPQTE	407
Db	406	fdrarkrigfavsachyvdefrtaavagpfylldmccgnvipote	452

[illegible]

Human aspartyl proteases can act as beta-secretase proteases useful for treating Alzheimer's disease. APP isoforms are useful for identifying modulators of amyloid-beta peptide production, for use in designing therapeutics for the treatment and prevention of Alzheimer's disease, dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis and neuronal loss. APP isoforms are also used in methods for identifying inhibitors and modulators of human Asp2 activity. The invention relates to a method for identifying agents that modulate the activity of human aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used as a means to screen in cellular assays for the inhibitors of beta- and gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in polymerase chain reactions (PCR). The probes are useful for detecting Hu-Asp nucleic acids in in vitro assays and in Northern and Southern blots. The present sequence is Human aspartyl protease 2a (Hu-Asp2a) deltaTM protein which is obtained by the deletion of transmembrane domain at the C-terminal end of Hu-Asp2a. Human Asp2a has beta-secretase activity.

Sequence 453 AA:

Query Match 100.0%; Score 2156; DB 22; Length 453;
Best Local Similarity 100.0%; Pred. No. 2,4e-210;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ETDEPEEPRGRSGSEVEMVDNLRGKSGGYVEMTVGSPPTLNLIVDTGSSNFAVGAAP 60
|||||
46 etdeepeepgrsgsfvemdnlrgksggyvemtvgspptlnllydvsgsnfavgaap 105
61 HPELHRYTORLSSTYRDLRKGYVPYTGKGWEGELGTLVSIIPGPNVYRANIAATTE 120
|||||
106 hpehlhrytorlssstyrdlrkgyvpytgkgwegelgtlvsipgpnvyraniaalte 165
121 SDFEFINGSNMGEILGLAYAEIARPDSDLPEPDSLVKOTHPNLEFSLQCGAGPELND 180
|||||
166 sdfefingsnmgeilglayaeiarpdsslpepdsivkqthvnpnlfsqlcgagpflnqs 225
181 EYLAASVGSMTIGIDHSLYTGLWTPTRREKYEVIIIVAREINODKMKCKENYDK 240
|||||
226 eylasvgsmtigidhslytglwtptrreyeviiivareinodkmkckenynydk 285
241 SYDSGEGTNLRLEPKYFEAAVKSIAASSTEKRPDGFMLGEOLVCMQAGTTPWNIFFVIS 300
|||||
286 sydsgegtnlrlrlepkylfeaaavksiaasstekrpdgfmvgelvcwagtltpwnilfpvis 345
301 LYLMGEVYNOSFRITLIPQOYLRPVEDVATSDCKKFAISOSSTGTWVGAVIMEGEYVY 360
|||||
346 lylmgevynosfriltipqoylrpvedvatstdckkfaissgstgtwvgavimegfyvv 405
361 FDRARRKRGFAVSAACHVDEFTAAVEGPFVTLDMEDCGYNIPTDE 407
|||||
406 fdarrkrigfavsachvdeftaaavegpfvtldmedcgynipptide 452

RESULT 5
AAU06616 ID AAU06616 standard; Protein; 453 AA.

XX AAU06616;

XX 24-OCT-2001 (first entry)

DE Human-pro-Asp 2(a) delta TM.

XX Human; Aspartyl protease; beta-secretase; noctropic; ASP2;

KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;

XX amyloid-beta; Abeta; Human-pro-Asp 2(a) delta TM, mutant; mutein.

OS Homo sapiens.

XX Synthetic.

FT Key

Location/Qualifiers
Peptide 1..21

FT Protein /label= signal_peptide
FT 22.453 /label= Mature_Human_Pro_Asp_2(a)_delta_TM
FT Misc-difference 214 /note= "Encoded by CAC"

PN WO200149098-A2.

XX 12-JUL-2001.

PF 09-MAY-2001; 2001WO-1B00798.

PR 09-MAY-2001; 2001WO-1B00798.

XX (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

XX (YANR/) YAN R.

PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;

XX WPI; 2001-502549/55.

DR N-PSDB; AAS11530.

XX Claim 149; Page 160; 185pp; English.

The invention relates to a purified polypeptide comprising a fragment of mammalian aspartyl protease (Asp2) protein which lacks the Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide and the fragment retain the beta-secretase activity of the mammalian Asp2 protein. The invention also details polynucleotides for the Asp2 protein and vectors expressing them, and a polypeptide (isoform of amyloid protein precursor (APP)) comprising the amino acid sequence of an APP or its fragment containing an APP cleavage site recognizable by a mammalian beta-secretase, and further comprising two lysine residues at the carboxyl terminus of the amino acid sequence of the mammalian APP or APP fragment. Also included in the invention are methods of identifying APP modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are useful for treating Alzheimer's disease. APP is useful in methods for identifying inhibitors or modulators of human Asp2 activity and amyloid-beta (Abeta) peptide production. APP is also useful in designing therapeutics for the treatment or prevention of Alzheimer's disease. APP comprising the APP-SW-beta-secretase peptide sequence (NLDA), which is associated with increased levels of Abeta processing is useful in assays relating the Alzheimer's research. The expression vector is useful for recombinantly expressing APP. Nucleic acids that hybridize to CC Asp oligonucleotides are useful as probes or primers. The probes are useful for detecting Hu-Asp nucleic acids in in vitro assays and in Northern and Southern blots. The present sequence is Human-pro-Asp 2(a) delta TM protein, which lacks the C-terminal transmembrane domain.

XX Sequence 453 AA:

Query Match 100.0%; Score 2156; DB 22; Length 453;
Best Local Similarity 100.0%; Pred. No. 2,4e-210;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 ETDEPEEPRGRSGSEVEMVDNLRGKSGGYVEMTVGSPPTLNLIVDTGSSNFAVGAAP 60

DB 46 etdeepeepgrsgsfvemdnlrgksggyvemtvgspptlnllydvsgsnfavgaap 105

XX 61 HPELHRYTORLSSTYRDLRKGYVPYTGKGWEGELGTLVSIIPGPNVYRANIAATTE 120

DB 106 hpehlhrytorlssstyrdlrkgyvpytgkgwegelgtlvsipgpnvyraniaalte 165


```

FH Key Location/Qualifiers
FT Misc-difference 214 /note= "Encoded by CAC"
XX
XX PN MO200123533-A2.
XX
XX PD 05-APR-2001.
XX
XX PE 22-SEP-2000; 2000WO-US26080.
XX
XX PR 23-SEP-1999; 99US-0155493.
XX PR 23-SEP-1999; 99WO-US20881.
XX PR 13-OCT-1999; 99US-0416901.
XX PR 06-DEC-1999; 99US-0169232.
XX
XX PA (PHRA ) PHARMACIA & UPJOHN CO.
XX
XX PI Gurney M, Bienkowski MJ;
XX
XX DR WPI; 2001-290516/30.
XX DR N-PSDB; AAD06752.
XX
XX PT Enzymes that cleave the alpha-secretase site of the amyloid precursor
XX PT protein, useful for the treatment of Alzheimer's disease -
XX
XX PS Example 10; Fig 11; 189pp; English.
XX
XX CC The present invention relates to enzymes for cleaving the alpha-
XX CC secretase site of the amyloid precursor protein (APP) and methods of
XX CC identifying those enzymes. The methods may be used to identify enzymes
XX CC that may be used to cleave the alpha-secretase cleavage site of the APP
XX CC protein. The enzymes may be used to treat or modulate the progress of
XX CC Alzheimer's disease. The present sequence is human Aspartyl protease 2a
XX CC (Asp-2a) deltatm protein which is obtained by deleting its transmembrane
XX CC domain. This sequence has beta-secretase protease activity.
XX
XX SQ Sequence 453 AA:

```

Query Match 100.0%; Score 2156; DB 22; Length 453;
Best Local Similarity 100.0%; Pred. No. 2.4e-210;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ENDEEPEEGRGRSEFVEMVDNLRGSGGQYVEMTVGSPPTLNILVDTGSSNFAVGAAP 60
DB 46 etdeepeepgrgrsfevmdnlrgsggqyvevmvgsppqlnlldvdtgssnfavgaap 105
QY 61 HPLHRYGROLSSYTRDLRKGVYVPTQGMKEGELGTDLVSIPIHGPNTVYRANIAATE 120
DB 106 hpllhrygrqlsstyrdlrkgyvpytqgmkegelgtdlvsiiphgpnvtraniaaate 165
QY 121 SDRFFINGSNWEGILGLAETARPPDLSLEPPFDSLVRQTHVNFSLQLCGAGFPLNOS 180
DB 166 sdrffingsnwegilglayaelarppdlslepfdsivkqthvnlfsqlcgagfplnos 225
QY 181 EYLASVGSNIIIGIDHSILYTGSLMTPTIRREYVEVITIRVINGODLKMCKEKNYDK 240
DB 226 eylasvgsniiigidhsilysglwtpirreyvevilitrvringodlkmckeknydk 285
QY 241 SIYDSGTNTLRPKKVFEEAIVKSIKAASSTKEKFPDGFMLGEOLVCMWAGTTPMNIPFVIS 300
DB 286 siydsgtntlrpkkvfeavvksikaasstekfpdgfwlgeqlvcmwagtlpmnifpvis 345
QY 301 LYLMGEVYNOSFRITLIPQOYLARVEDVATSDCCYKFAISOSSTGVNMGAVIMEGYVYV 360
DB 346 lylmgevynosfrtllipqyylrpvedvatsdcdcykfaitsgstgvnmgavimegylyv 405
QY 361 FDRARRKILGFAVSACHVDEFRFAAVEGPVTLDMEDCGYNIPTQDE 407
DB 406 fdarrkiligfavsachvdefrfaavegpvtlmedcgyniptqde 452

```

```

AAB07897
ID AAB07897 standard; Protein; 456 AA.
XX
XX AC AAB07897;
XX
XX DT 14-NOV-2000 (first entry)
XX
XX DE Active enzyme portion of human beta-secretase enzyme.
XX
XX KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
XX KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
XX KW inhibitor; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200047618-A2.
XX
XX PD 17-AUG-2000.
XX
XX PE 10-FEB-2000; 2000WO-US03819.
XX
XX PR 10-FEB-1999; 99US-0119571.
XX PR 15-JUN-1999; 99US-0139172.
XX
XX PA (ELAN-) ELAN PHARM INC.
XX
XX PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
XX PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX
XX DR WPI; 2000-533011/48.
XX
XX PT Purified beta-secretase protein used in assays to discover inhibitors
XX PT which can be used for the treatment of amyloidogenic diseases e.g.
XX PT Alzheimer's disease -
XX
XX PS Claim 24; Fig 2b; 121pp; English.
XX
XX CC The specification describes a beta-secretase enzyme. The enzyme cleaves
XX CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
XX CC enzyme is therefore implicated in the production of amyloid plaque
XX CC components which accumulate in the brains of individuals afflicted with
XX CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
XX CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
XX CC disease-like pathology to test if they maintain or improve cognitive
XX CC ability or reduce the plaque burden. The compounds are used for the
XX CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
XX CC present sequence represents the active enzyme portion of human
XX CC beta-secretase enzyme.
XX
XX SQ Sequence 456 AA:

```

Query Match 100.0%; Score 2156; DB 21; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.4e-210;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ENDEEPEEGRGRSEFVEMVDNLRGSGGQYVEMTVGSPPTLNILVDTGSSNFAVGAAP 60
DB 1 etdeepeepgrgrsfevmdnlrgsggqyvevmvgsppqlnlldvdtgssnfavgaap 60
QY 61 HPLHRYGROLSSYTRDLRKGVYVPTQGMKEGELGTDLVSIPIHGPNTVYRANIAATE 120
DB 61 hpllhrygrqlsstyrdlrkgyvpytqgmkegelgtdlvsiiphgpnvtraniaaate 120
QY 121 SDRFFINGSNWEGILGLAETARPPDLSLEPPFDSLVRQTHVNFSLQLCGAGFPLNOS 180
DB 121 sdrffingsnwegilglayaelarppdlslepfdsivkqthvnlfsqlcgagfplnos 180
QY 181 EYLASVGSNIIIGIDHSILYTGSLMTPTIRREYVEVITIRVINGODLKMCKEKNYDK 240
DB 181 eylasvgsniiigidhsilysglwtpirreyvevilitrvringodlkmckeknydk 240
QY 241 SIYDSGTNTLRPKKVFEEAIVKSIKAASSTKEKFPDGFMLGEOLVCMWAGTTPMNIPFVIS 300

```


modulators of amyloid-beta peptide production, for use in designing
therapeutics for the treatment and prevention of Alzheimer's disease,
dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
and neuronal loss. APP isoforms are also used in methods for identifying
inhibitors and modulators of human Asp2 activity. The invention relates
to a method for identifying agents that modulate the activity of human
aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
as a means to screen in cellular assays for the inhibitors of beta- and
gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
polymerase chain reactions (PCR). The probes are useful for detecting
Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
blots. The present sequence is Human aspartyl protease 2a (Hu-Asp2a)
deltaTM (His)6 protein which is obtained by deletion of C-terminal
transmembrane domain and addition of a hexa-Histidine tag at the
C-terminal end of Hu-Asp2a. Human Asp2a has beta-secretase activity.

FT /label= Mature_Human_Pro-Asp-2(a)_delta-TM_(His)6
FT Misc-difference 214 /note= "Encoded by CAC"
FT Misc-difference 454..459
FT Region /note= "Encoded by CAGCAGCAGCAGCAGCAGC"
FT 454..459
FT /label= His tag
FT /note= "Nickel binding region to aid purification"
PN WO200149098-A2.
PD 12-JUL-2001.
XX
PD 09-MAY-2001; 2001WO-IB00798.
XX
PF 09-MAY-2001; 2001WO-IB00798.
XX
PR 09-MAY-2001; 2001WO-IB00798.
XX
PA (BIEN/) BIENKOWSKI M J.
XX (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
XX (PARO/) PARODI L A.
PA (YANR/) YAN R.

PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;
XX WPI; 2001-502549/55.
XX N-PSDB; AAS11531.

PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PS activity -
XX Claim 149; Fig 12; 185pp; English.

XX The invention relates to a purified polypeptide comprising a fragment of
CC mammalian aspartyl protease (Asp)2 protein which lacks the Asp2
CC transmembrane domain and the Asp2 protein, and where the polypeptide and
CC the fragment retain the beta-secretase activity of the mammalian Asp2
CC protein. The invention also details polynucleotides for the Asp
CC proteins and vectors expressing them, and a polypeptide (isoform of
CC amyloid protein precursor (APP)) comprising the amino acid sequence of an
CC APP or its fragment containing an APP cleavage site recognizable by a
CC mammalian beta-secretase, and further comprising two lysine residues at
CC the carboxyl terminus of the amino acid sequence of the mammalian APP or
CC APP fragment. Also included in the invention are methods of identifying
CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
CC useful for treating Alzheimer's disease. APP is also useful in methods for
CC identifying inhibitors or modulators of human Asp2 activity and
CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
CC APP comprising the APP-Sw-beta-secretase peptide sequence (NDA), which
CC is associated with increased levels of Abeta processing is useful in
CC assays relating the Alzheimer's research. The expression vector is useful
CC for recombinantly expressing APP. Nucleic acids that hybridize to
CC Asp oligonucleotides are useful as probes or primers. The probes are
CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC Northern and Southern blots. The present sequence is Human-pro-
CC Asp 2(a) delta TM (His)6 protein, which lacks the C-terminal
CC transmembrane domain and has a His tag to aid purification.
XX

Sequence 459 AA:

Query Match 100.0%; Score 2156; DB 22; Length 459;
Best Local Similarity 100.0%; Pred. No. 2.5e-210; Indels 0; Gaps 0;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEEPEEPGRSGFVEMVDNLGRKSGGGYVEMTVGSPPTQNTIIVDGGSSNFAGAAP 60
DB 46 etdeepeepgrsgfivemvdnlgrksgggyvemtvgspptqntilivdggssnfavaap 105
QY 61 HPLHRYRQQLSTYRDLRKGYVYPTGCKWEGELGTDLVSIIPHGPNTVRANIAITE 120

Query Match 100.0%; Score 2156; DB 22; Length 459;
Best Local Similarity 100.0%; Pred. No. 2.5e-210; Indels 0; Gaps 0;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETDEEPEEPGRSGFVEMVDNLGRKSGGGYVEMTVGSPPTQNTIIVDGGSSNFAGAAP 60
DB 46 etdeepeepgrsgfivemvdnlgrksgggyvemtvgspptqntilivdggssnfavaap 105
QY 61 HPLHRYRQQLSTYRDLRKGYVYPTGCKWEGELGTDLVSIIPHGPNTVRANIAITE 120
DB 106 hplhryrqqlstyrdlrkgyvpytgkwegelgtdlvsiiphgpntrvaniaite 165
QY 121 SDFEFGNSWEGILGAAVAETARPDDSLPEPDSLVTQTHVNFSLQCGAGFPLNOS 180
DB 166 sdfefngswegilgaavaeetarppddslpepdslyvtqhvnfslqcgagfplngs 225
QY 181 EVLASVGGSMIIIGIDHSYTGSLMTPIRREWEYEVIIIVREINSGODKMKCKENYVK 240
DB 226 evlasvggsmiigidhslytgslytprreweyeviiivreiingodkkmckekenyk 285
QY 241 SYDSTTNRLPKKRYFEAAVSKRAASSTEFKPDGFMIGEDLVCMQAGTTPWNIFFPVIS 300
DB 286 sydsttnrlpkkryfeaaavskraasstekfpdgfmigedlvcmwagttcpwnifpvls 345
QY 301 LYLMEGVTNQSFRTITLPOQYLRPEVDVATSDDDCKKFAISQSSSTGTWGVAIMGFFYV 360
DB 346 lylmevtnqsfritlppqylrpvedvatsddckkfaissstgtwgvaimgffyvv 405
QY 361 FBRARRRIGFAVSACHVDEFTAAVEGPFVTLDMEDCGYNIPQTD 407
DB 406 fbrarrkigfavsachvdeftaaavgpfvltlmedcgynlptqde 452
RESULT 12
AAU06617
ID AAU06617 standard; Protein: 459 AA.
XX
AC AAU06617;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human-pro-Asp 2(a) delta TM (His)6.
XX
KW Human; Aspartyl protease; beta-secretase; nootropic; Asp2;
KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
XX amyloid-beta; Abeta; Human-pro-Asp 2(a) delta TM (His)6; mutant; mutlein.
OS Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= signal_peptide
FT Protein 22..459

ID AAE02595 standard; Protein: 459 AA.
XX
AC AAE02595;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human-Asp-2(a) deltatm (His)6 protein.
XX
KW Human; alpha-secretase; amyloid precursor protein; APP; therapy;
KW Alzheimer's disease; antialzheimer's; aspartyl protease 2a; Asp2a;
KW beta-secretase; Asp-2a delta tm; histidine tag; mutant; muten.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 214 /note= "Encoded by CAC"
FT MISC-difference 454 /note= "Encoded by CAC"
FT MISC-difference 455 /note= "Encoded by CAC"
FT MISC-difference 456 /note= "Encoded by CAC"
FT MISC-difference 457 /note= "Encoded by CAC"
FT MISC-difference 458 /note= "Encoded by CAC"
FT MISC-difference 459 /note= "Encoded by CAC"
FT MISC-difference 459 /note= "Encoded by CAC"
FT MISC-difference 459 /note= "Encoded by CAC"
XX
PN WO200123533-A2.
XX
PD 05-APR-2001.
XX
PE 22-SEP-2000; 2000WO-US26080.
XX
PR 23-SEP-1999; 99US-0155493.
PR 23-SEP-1999; 99WO-US20881.
PR 13-OCT-1999; 99US-0416901.
PR 06-DEC-1999; 99US-0169232.
XX
PA (PHMA) PHARMACIA & UPJOHN CO.
XX
PI Gurney M, Blenkowski MJ;
XX
DR MPI: 2001-290516/30.
DR N-PSDB; AAD06753.
XX
PT Enzymes that cleave the alpha-secretase site of the amyloid precursor
PT protein, useful for the treatment of Alzheimer's disease -
XX
PS Example 10; Fig 12; 189pp; English.
XX
CC The present invention relates to enzymes for cleaving the alpha-
CC secretase site of the amyloid precursor protein (APP) and methods of
CC identifying those enzymes. The methods may be used to identify enzymes
CC that may be used to cleave the alpha-secretase cleavage site of the APP
CC protein. The enzymes may be used to treat or modulate the progress of
CC Alzheimer's disease. The present sequence is human Aspartyl protease 2a
CC (Aap 2a) delatm (His)6 protein which is obtained by deleting the
CC transmembrane domain and adding a histidine tag at the C-terminal end.
CC This sequence has beta-secretase protease activity.
XX
XX
SO Sequence 459 AA;

Query Match 100.0%; Score 2156; DB 22; Length 459;
Best Local Similarity 100.0%; Pred. No. 2.5e-210;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEPEEPGRGSPVEVDLRRKSGOGYVETVGSPPQTLIIIVDTGSSNFVAGAP 60
|||||

Db 46 etdepeepgrigsfvevndljgksqggyvemtvtgspptlnllvdtgssnfavgaap 105
QY 61 HPELHRYGROLSSRYRDLRKGVYVETVQKMEGELGPTDNLSPHGPNTVANAIAATE 120
Db 106 hpfclhryqqlaslyrdlrkgyvpytqgkwegelgtlvsilphgpnvtvraanaaite 165
QY 121 SDKFFINGSNWBSITGLAVAEIARPDSDLPPFDSLVRQTHPNFLSLOLCAGFPLNOS 180
Db 166 sckffingsnwegllglayaelarpdsslppfdslyvqthvnpnlfsiqgagfplpns 225
QY 181 EYLVASVGSMTIIGGIDHSIYTGSLWTPYRREYVEVLIIVRVEINGODIKMDCKEYNDK 240
Db 226 evlvasvgsmtiigglidhslytgslywtpyrrewyevliivrvellingd1kmdckeyndk 285
QY 241 SIYDSCGTNLRPLPKVFEAAVKSIRKASTSEKFPFGFNLGEOYQWOGTTPWNTFFPIS 300
Db 286 slvdsqgtlnlrplpkvfeavksirkaastekfpqgfalgeqlvewqgqtlpwnlfpvis 345
QY 301 LYIMGEVTVNQSFRTITLPOQYLRPEVDVATSDDCYKFAISQSSSTGYVAGAVIMEGFYV 360
Db 346 lylmgevtnqsfrtltlpqylrpvedvatsqdcyckfaissstgvtmgavimegfyyv 405
QY 361 PDRARRIGFAVSACHVDEPFAVEGPPVTLMDGCGYNIPQTD 407
Db 406 fdarrkrigfavsachvdefptaavegppvtlmdcgynipqtde 452

RESULT 15
AAB07898
ID AAB07898 standard; Protein: 460 AA.
XX
AC AAB07898;
XX
DT 14-NOV-2000 (first entry)
XX
DE Amino acid sequence of a human beta-secretase enzyme fragment.
XX
KW Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor.
XX
OS Homo sapiens.
XX
OS
XX
PN WO200047618-A2.
XX
PD 17-AUG-2000.
XX
PE 10-FEB-2000; 2000WO-US03819.
XX
PR 10-FEB-1999; 99US-0119571.
PR 15-JUN-1999; 99US-0139172.
XX
PA (ELAN-) ELAN PHARM INC.
XX
PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Slnha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX
DR MPI: 2000-533011/48.
XX
PT Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease -
XX
XX
PS Claim 55; Fig 3A; 121pp; English.
XX
CC The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the

CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a human beta-secretase enzyme fragment.
XX
SQ Sequence 460 AA;

Query Match 100.0%; Score 2156; DB 21; Length 460;
Best Local Similarity 100.0%; Pred. No. 2.5e-210;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEPEEPGRKSGFEVMDNLKSKSQGYVEMTVSGSPQTLNLLVDTGSSNFAYGAAP 60
DB 46 etdepeepgrksgfevmdnlrksqgyvemtvsppqlnllvdtgssnfavgaap 105
QY 61 HPELHRYQRQLSTYRDLRKGVVVPYTOGKMEGELGTDLVSIPHGPNTVRANIAITE 120
DB 106 hpeLhryqrqlstYrDLrKGVVVPYtOGkMEgELgTDlVsIpHGPnTVrAnIAiTe 165
QY 121 SDKFFINGSNMEGILGLAYAEIARPPDLSLEPFSDLYKQTHVPMLESLQCGAGFPLNOS 180
DB 166 sdkffingsnmeGIlglAYAEIARpPDLsLEpFSDlYkQThVPMLEsLQcGAGfPlNOS 225
QY 181 EVLASVGSMLIGGDHSLYTGSLMTPIRREWYEVIIVREINGODLKMCKEYNDK 240
DB 226 evlasvgSMLIGGDhSLYtGSLMTPIRrEWYEVIIvREInGODlKMcKEyNDK 285
QY 241 STVDSGTNLRLPKKVFEEAAVKSIRAASTERKFPDGFMLGEOLYCWOAGTTPWNIFFYIS 300
DB 286 stvdsgtNLRlPKKvFEeAAVKSIRAAStERkFPDGFmLGEOlYcWOAGtTPWNIffYIS 345
QY 301 IYLMGEVTNQSFRITILLPOQYLRPEVDVATSQDDCYKFAISQSTGYMGAIVMEGFYV 360
DB 346 iYlMgevtNqSFRITIlLpOqYLRpEvdVATsQDDCYkFAISqSTGyMGAIVMEGFyV 405
QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGFVTLDMEDCGYNIPQTD 407
DB 406 fdarrrRiGfAVsACHvHdEfRtAAvEGfVtLdMEdCGyNIpQtd 452

Search completed: August 7, 2002, 09:16:15
Job time: 160 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2002, 09:14:48 ; Search time 22.58 seconds

(without alignments)
440.266 Million cell updates/sec

Title: US-09-724-571-58

Perfect score: 2156

Sequence: 1 ERDEPEREGRGRGSEVEMD.....GEFTLMDGCGYNIPQDE 407

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2150	99.7	501	US-09-009-191-2	Sequence 2, Appl1
2	2073	96.2	774	US-09-009-191-4	Sequence 2, Appl1
3	1132	52.5	518	US-08-999-723-2	Sequence 2, Appl1
4	1132	52.5	518	US-09-434-427-2	Sequence 2, Appl1
5	1116.5	51.8	514	US-09-717-432-2	Sequence 2, Appl1
6	1116.5	51.8	514	US-09-913-484-2	Sequence 2, Appl1
7	299.5	13.9	396	US-08-208-007A-13	Sequence 13, Appl1
8	299.5	13.9	396	US-09-032-523-9	Sequence 9, Appl1
9	293.5	13.7	412	US-08-208-007A-12	Sequence 12, Appl1
10	293.5	13.7	412	US-08-974-691-4	Sequence 4, Appl1
11	279.5	13.0	458	US-09-640-305-6	Sequence 6, Appl1
12	273.5	12.7	409	US-08-846-021A-8	Sequence 8, Appl1
13	273.5	12.7	409	US-08-088-633-2	Sequence 2, Appl1
14	273.5	12.7	427	US-08-441-751-2	Sequence 2, Appl1
15	270	12.5	410	US-08-245-756-2	Sequence 2, Appl1
16	270	12.5	410	US-08-441-751-2	Sequence 2, Appl1
17	270	12.5	410	US-08-441-751-2	Sequence 2, Appl1
18	270	12.5	410	US-08-441-751-2	Sequence 2, Appl1
19	270	12.5	410	US-08-441-751-2	Sequence 2, Appl1
20	250	11.6	349	PCT-US92-02521-2	Sequence 2, Appl1
21	227	10.5	398	US-09-032-523-3	Sequence 2, Appl1
22	227	10.5	398	US-08-328-314-2	Sequence 2, Appl1
23	212	9.8	419	US-08-974-691-3	Sequence 2, Appl1
24	210	9.7	397	US-09-079-415-2	Sequence 2, Appl1
25	196.5	9.1	430	US-08-535-237-2	Sequence 2, Appl1
26	195.5	9.1	427	US-07-958-222A-2	Sequence 2, Appl1
27	193	9.0	330	US-08-115-753-1	Sequence 1, Appl1

28	193	9.0	419	US-08-115-753-2	Sequence 2, Appl1
29	193	9.0	419	US-08-115-753-3	Sequence 33, Appl1
30	187	8.7	420	US-09-008-271A-4	Sequence 4, Appl1
31	187	8.7	420	US-08-974-691-8	Sequence 8, Appl1
32	184.5	8.6	395	US-08-723-938-3	Sequence 3, Appl1
33	184.5	8.6	395	US-09-080-538-3	Sequence 3, Appl1
34	184	8.5	445	US-08-974-691-6	Sequence 6, Appl1
35	184	8.5	451	US-08-974-691-2	Sequence 2, Appl1
36	149	6.9	437	US-09-353-332-2	Sequence 2, Appl1
37	129.5	6.0	140	US-09-211-631-13	Sequence 13, Appl1
38	129.5	6.0	140	US-09-265-628-13	Sequence 13, Appl1
39	129.5	6.0	140	US-09-001-141-11	Sequence 11, Appl1
40	129.5	6.0	140	US-09-532-803-6	Sequence 6, Appl1
41	129.5	6.0	140	US-09-653-403-14	Sequence 14, Appl1
42	97	4.5	1030	US-09-091-117-2	Sequence 2, Appl1
43	95.5	4.4	280	US-09-160-246-14	Sequence 14, Appl1
44	89.5	4.2	1097	US-08-680-326-39	Sequence 39, Appl1
45	88	4.1	746	US-08-838-219B-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-009-191-2
; Sequence 2, Application US/09009191
; Patent No. 63196689
; GENERAL INFORMATION:
; APPLICANT: POWELL, DAVID
; APPLICANT: CHAPMAN, CONRAD
; APPLICANT: MURPHY, KAY
; APPLICANT: SMITH, TRUDI
; TITLE OF INVENTION: ASP2
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESS: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,191
; FILING DATE: 20-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 9701684.4
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70368
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-009-191-2
Query Match 99.7%; Score 2150; DB 4; Length 501;
Best Local Similarity 99.8%; Pred. No. 9,4e-223;

Matches 406; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETDEPEPEGRGSGFVEMVNDLRKSGGGYVEMTVGSPPTQTLNLTVDGSSNFVGAAP 60
 Db 46 ETDEPEPEGRGSGFVEMVNDLRKSGGGYVEMTVGSPPTQTLNLTVDGSSNFVGAAP 105

QY 61 HPFLHRYQRLSSTYRDLRKGVVVPYTOGKMEGELGTDLSIPHGPNVYRANIAITE 120
 Db 106 HPFLHRYQRLSSTYRDLRKGVVVPYTOGKMEGELGTDLSIPHGPNVYRANIAITE 165

QY 121 SDKFFINGSNEGILGLAYAEIARPDSDLPEPFDLSVKQTHVPLNLSIQLCGAGFPLNOS 180
 Db 166 SDKFFINGSNEGILGLAYAEIARPDSDLPEPFDLSVKQTHVPLNLSIQLCGAGFPLNOS 225

QY 181 EYLASVSGSMIIGIDHSLYTGSLMTPTIRREMYEVIIIVVEINGODLKMCKEYNDK 240
 Db 226 EYLASVSGSMIIGIDHSLYTGSLMTPTIRREMYEVIIIVVEINGODLKMCKEYNDK 285

QY 241 STVDSGTTNLRPKKVFEEAAVKSIAASSTKFPDGFMLGEOLVCWQAGTTPMNIFFVIS 300
 Db 286 STVDSGTTNLRPKKVFEEAAVKSIAASSTKFPDGFMLGEOLVCWQAGTTPMNIFFVIS 345

QY 301 IYLMGEVYNOSRTITLPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 360
 Db 346 IYLMGEVYNOSRTITLPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 405

QY 361 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDE 407
 Db 406 FDRARRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDE 452

RESULT 2
 US-09-009-191-4
 ; Sequence 4, Application US/09009191
 ; Patent No. 6319689
 ; GENERAL INFORMATION:
 ; APPLICANT: POWELL, DAVID
 ; APPLICANT: CHAPMAN, CONRAD
 ; APPLICANT: MORPHY, KAY
 ; APPLICANT: SMITH, TRUDI
 ; TITLE OF INVENTION: ASP2
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: RATNER & PRESTIA
 ; STREET: P.O. BOX 980
 ; CITY: VALLEY FORGE
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19482
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/009,191
 ; FILING DATE: 20-JAN-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: UK 9701684.4
 ; FILING DATE: 28-JAN-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PRESTIA, PAUL F
 ; REGISTRATION NUMBER: 23,031
 ; REFERENCE/DOCKET NUMBER: GH-70368
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-407-0700
 ; TELEFAX: 610-407-0701
 ; TELEX: 846169
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 774 amino acids
 ; TYPE: amino acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-009-191-4

Query Match 96.2%; Score 2073; DB 4; Length 774;
 Best Local Similarity 99.2%; Pred. No. 3,8e-214;
 Matches 392; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 13 GSFVEVNDLRKSGGGYVEMTVGSPPTQTLNLTVDGSSNFVGAAPFLHRYQRL 72
 Db 1 GSFVEVNDLRKSGGGYVEMTVGSPPTQTLNLTVDGSSNFVGAAPFLHRYQRL 60

QY 73 SSTYRDLRKGVVVPYTOGKMEGELGTDLSIPHGPNVYRANIAITE SDKFFINGSNME 132
 Db 61 SSTYRDLRKGVVVPYTOGKMEGELGTDLSIPHGPNVYRANIAITE SDKFFINGSNME 120

QY 133 GILGLAYAEIARPDSDLPEPFDLSVKQTHVPLNLSIQLCGAGFPLNOS EYLASVSGSMII 192
 Db 121 GILGLAYAEIARPDSDLPEPFDLSVKQTHVPLNLSIQLCGAGFPLNOS EYLASVSGSMII 180

QY 193 GGDHSLYTGSLMTPTIRREMYEVIIIVVEINGODLKMCKEYNDKSI VDSGTTNLR 252
 Db 181 GGDHSLYTGSLMTPTIRREMYEVIIIVVEINGODLKMCKEYNDKSI VDSGTTNLR 240

QY 253 PKKVFEEAAVKSIAASSTKFPDGFMLGEOLVCWQAGTTPMNIFFVIS IYLMGEVYNOSF 312
 Db 241 PKKVFEEAAVKSIAASSTKFPDGFMLGEOLVCWQAGTTPMNIFFVIS IYLMGEVYNOSF 300

QY 313 RITILPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV FDRARRIGFAV 372
 Db 301 RITILPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV FDRARRIGFAV 360

QY 373 SACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDE 407
 Db 361 SACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQDE 395

RESULT 3
 US-08-999-723-2
 ; Sequence 2, Application US/08999723A
 ; Patent No. 6025180
 ; GENERAL INFORMATION:
 ; APPLICANT: Powell, David J.
 ; APPLICANT: Southan, Christopher
 ; APPLICANT: Chapman, Conrad G.
 ; APPLICANT: Evans, Joanne R.
 ; TITLE OF INVENTION: ASP1
 ; FILE REFERENCE: GH70262
 ; CURRENT APPLICATION NUMBER: US/08/999,723A
 ; CURRENT FILING DATE: 1997-10-06
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patencin Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 518
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-08-999-723-2

Query Match 52.5%; Score 1132; DB 3; Length 518;
 Best Local Similarity 54.0%; Pred. No. 4.1e-113;
 Matches 209; Conservative 65; Mismatches 109; Indels 4; Gaps 2;

QY 9 PCRGRSGFVEMVNDLRKSGGGYVEMTVGSPPTQTLNLTVDGSSNFVGAAPFLHRY 68
 Db 71 PAGAANFLAMVNDLIGSGRGLYLEMIGPPQKQIILVDGSSNFVGAAGTTPHSTIDYF 130

QY 69 QRLSSTYRDLRKGVVVPYTOGKMEGELGTDLSIPHGPNVYRANIAITE SDKFFING 128
 Db 131 DTERSSYTRKSGFDVTVYKQSGMTGFGEDLVITIKGFNTSLVAINATIFSENFELPG 190

[illegible]

```

QY      8  EPER-NGSPFVAVNDLRLKRSOQGYVVEKTVSSPPQTLNLIYDTGSSNFVGAAPHPELHR 66
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      65  EPKATANFLAVNDLQDSDGSGYILEMLIGTPQKVOILLVDGSSNFVGAAPHSHYIDT 122
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      67  YXOROLSTYRDLRGKGVVYPYTOGMEGELGTDLVSIPIHPNVTYVANIATTESDKFFI 122
      ||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      125  YFDESSESTYHSKGDFVYKTYTOGSMWGFVGEDLVYTIKPFNFSSFLVNIATITFESENFPL 182
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      127  NGSNMGILIGLAKAIARPDDSLPEFPDLSVKQTHPNPLFSLQLCAGFPLNQSEVLASY 180
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      185  PGKMGWILIGLAVAAALARPSSLETFPDSLVQAQKIPDJFSMQMGAGLPVAGS---GTN 242
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      187  GSGMIIGIDHSLYGLSWMTYPIRREWYVEYIIVREINODLKMCKREYNDKSTVSG 244
      ||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      242  GSGVLVIGIEPSLYIGDWTYPIKEEMTYQTEIUKETIGSONLNDCKRYNADKRAIVSG 307
      ||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      247  TTLNLRPKKVFEEAAVKSIRKAASSTREKPPDFWGLGEOLVYCAQTTPWNIIPVLSYLNGE 304
      ||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      302  TTLNLRPKQKVFADVAEAVARTSLIPEPSDGFWMGAQLACTENSEIPMAFPKISTYLRDE 366
      ||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      307  VNQSRRTIILPQQLRVREVEVAISQDDCKKFAISQSSSTGYWGAIVMEGYVVFDRARK 366
      :::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      362  NMSRSRTIILPOLYIQPMGAGFNY-ECYRFGISSSTALVIGATVMEGFEVVEDRQOR 420
      :::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY      367  RIGFAVASCHEDEFRTAAVCGPEVTLDM 395
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      421  RVGFAPVSPCAIEIGTVYSEISGPFSTEDI 449
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT      6
US-09-912-484-2
: Sequence 2, Application us/09912484
: Patent No. 6358725
: GENERAL INFORMATION:
: APPLICANT: Christie, Gary
: APPLICANT: Li, Xiaotong
: APPLICANT: Powell, David J.
: APPLICANT: Zhu, Yuan

```

TITLE OF INVENTION: Mouse Aspartic Secretase-1 (masp1)
 FILE REFERENCE: GP-70663-D1
 CURRENT APPLICATION NUMBER: US/09/912,484
 CURRENT FILING DATE: 2001-07-25
 PRIOR APPLICATION NUMBER: 60/166,974
 PRIOR FILING DATE: 1999-11-23
 PRIOR APPLICATION NUMBER: 09/717,432
 PRIOR FILING DATE: 2000-11-21
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 2
 LENGTH: 514
 TYPE: PRT
 ORGANISM: MUS MUSCULUS
 US-09-912-484-2

Query Match 51.8%; Score 1116.5; DB 4; Length 514;
 Best Local Similarity 53.7%; Pred. No. 1.9e-111;
 Matches 209; Conservative 66; Mismatches 109; Indels 5; Gaps 3;

8 EPGR-RGSFVEVADNLKRGSGGYVEMTVGSPPTQTLNLIVDTGSSNFAVGAAPPELHR 66
 65 EYRATANTLANVADNLQDGSRGYILEMLIGTPQKQILIVDTGSSNFAVGAAPHSYIDT 124
 67 YQRQLSTYRDLKRGVYPTQGMKEGLTDLYSIPHPNVTVRANITAEEDKPEI 126
 125 YDSSSSSYHSKGFVDVYKTYGSMTGFGVGDVLTIRKGFNSFLVNTATFESENFLL 184
 127 NCSNMEGILGLVAEIAEPDPSLEPFDSLVKQTHVNPILFSLQCGAGFPLNQEVLASV 186
 185 PEIKNGILGLVAAALAKSSLEPFDSLVQAOKIPDIFSMQCGAGLPVAGS--GTN 241
 187 GSGMIIGIDHSLYGLSLMTPIRREMYEVLIVREINGQDLKMDCKRYNDKSYDSG 246
 242 GCSLVIGIEPSLYKGDIVTYPIKEMYYQIELKLEIGOMINLDCREYNADKALVDSG 301
 247 TTNLRPKVFEAAVKSIAASTEKEPPDGFMLGEOLVCMQAGTPMNFPIYSILMGE 306
 302 TLLRTPQVFAVAVARTSLIPEFSDGFMTGAOLACTSETPMAFPKISYLRBE 361
 307 VNOSFRITLPOQLRPVEDVATSDDCYKFAISOSTGTGVAGVIMGEFYVFDRAK 366
 362 NMSRFRITLPOQLYIQPMGAGFNY-ECYRFGISSTNALVIGAVMEGFYVFDRAK 420
 367 RIGFAVSACHVDEFTAAVEGFVTLDM 395
 421 RVGFVAVSPCAIEGTVSEISGPFSTEDI 449

RESULT 7
 US-08-208-007A-13
 Sequence 13, Application US/08208007A
 Patent No. 5501969
 GENERAL INFORMATION:
 APPLICANT: HASTINGS, ET AL.
 TITLE OF INVENTION: Human Osteoclast-Derived Cathepsin
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARBILA, BYRNE, BAIN, GILFILLAN,
 ADDRESSEE: CECCHI, STEWART & OLSTEIN
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/208,007A

FILING DATE: March 8, 1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: NO. 5501969e
 FILING DATE: NO. 5501969e
 ATTORNEY/AGENT INFORMATION:
 NAME: FERRARO, GREGORY D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 325800-95
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 396 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-208-007A-13

Query Match 13.9%; Score 299.5; DB 1; Length 396;
 Best Local Similarity 25.9%; Pred. No. 1.2e-23;
 Matches 100; Conservative 65; Mismatches 148; Indels 73; Gaps 16;

3 DEEPEPGRGSEVMEVADNLKRGSGGYVEMTVGSPPTQTLNLIVDTGSSNFAVGA---- 58
 63 DOSAKP-----LNYLD-----MEYFGTISIGSPQNTVFIEDTGSSMLWPSVYCT 110
 59 AHPFLHRYQRLSTYRDLKRGVYPTQGMKEGLTDLYSIPHPNVTVRANITAI 118
 111 SPACKTHSRFOPSOSSSTYSOPQOSFSIQYGTGSLGIIGAQOVSV-EGITVVGQDFGESV 169
 119 TESDFEINGNWECHILGLVAEIAEPDPSLEPFDSLVKQTHVNPILFSLQCGAGFPLN 178
 170 TEPQGFVD-AEFDLILGLIGFSLA--VGATPYVDNMAQ-----NLVDLPMFSYMSN 222
 179 QSEVLASVGGSMIIGIDHSLYGLSLMTPIRREMYEVLIVREINGQDLKMDCKRYN 238
 223 PE---GGAGSELIIFGDIHSHSGSLNWPVKAWQIADLNIOVG--TFMFCSE--G 275
 239 DKSIVDSGTTNLRPKRYFEAAVKSIAASTEKEPPDGFMLGEOLVCMQAGTPMNFPIV 298
 276 COAIVDTGTSILTPGSDKIKQLQNAIGAP-----VDEYAVE-----CANLVMPD 322
 299 ISLYLMGEVTNOSFRITLPOQLRPVEDVATSDDCYKFAISOSTG----- 346
 323 VFTTNG-----VPTTSLPTAV--TLDVFDGMQFC-----SSGFGGLDIHPAPG 366
 347 -TVMGAVIMEGFYVFDRAKRIGFA 371
 367 LMTLDGVPIRQFYVFPDRGNRRVGLA 392

RESULT 8
 US-09-032-523-9
 Sequence 9, Application US/09032523
 Patent No. 6232454
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Corley, Neil C.
 APPLICANT: Guegler, Karl
 APPLICANT: Baugh, Mariah
 TITLE OF INVENTION: HUMAN PROTEINASE MOLECULES
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA

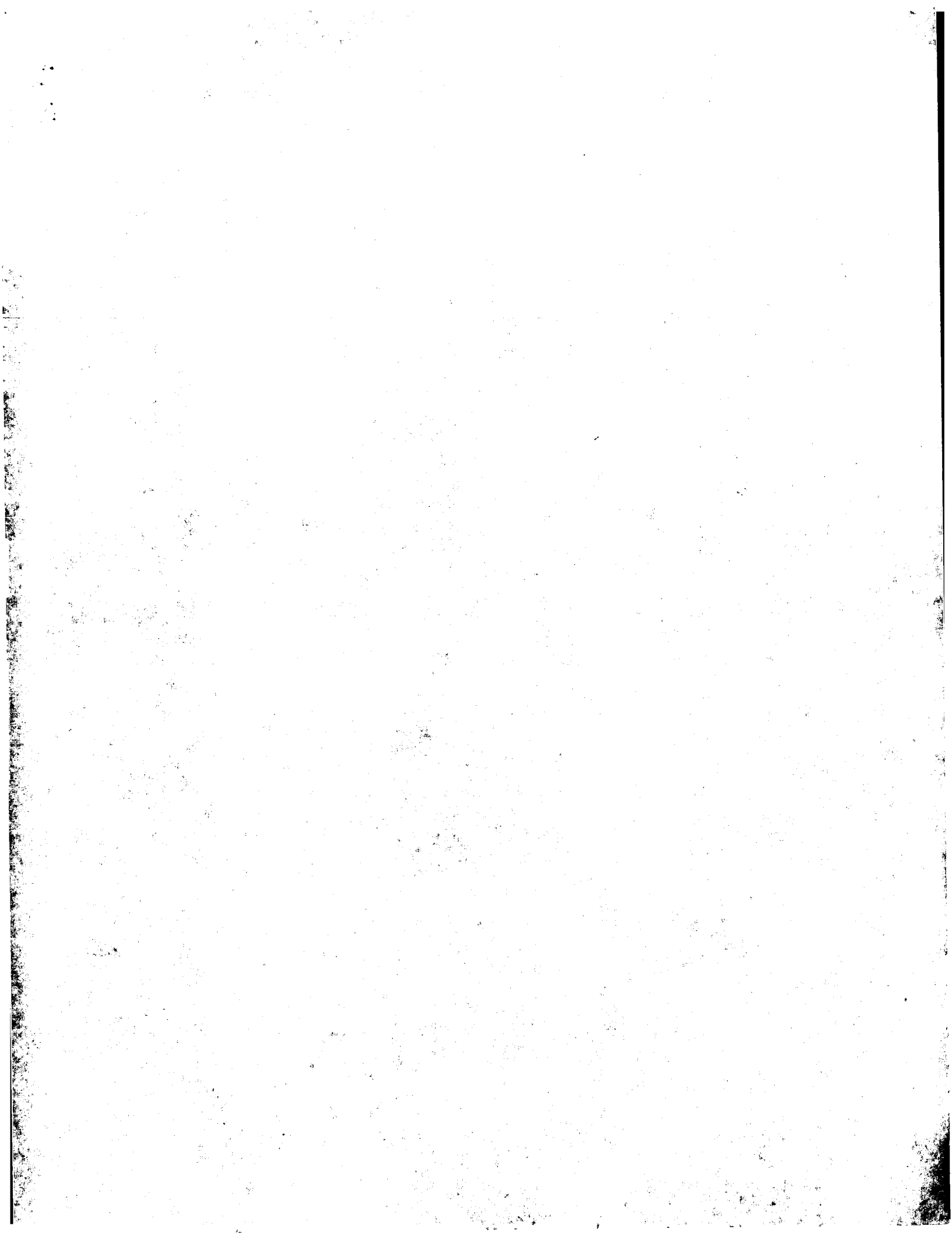
Wed Oct 30 14:11:51 2002

us-09-724-571-58.ra1

Page 9

DB 357 DYTEVSGSCISAFPMDFP-EPISPLAIGDSFLRKYYVYDLGDAVGLAKS 409

Search completed: August 7, 2002, 09:14:49
Job time: 74 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2002, 09:14:17 ; Search time 32.31 Seconds

(without alignments)
1210,410 Million cell updates/sec

Title: US-09-724-571-58

Perfect score: 2156

Sequence: 1 ETDEPEPEGRGSGFYEMVD.....GPFVTLDMEDCGYNIPQTE 407

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2156	100.0	501	2	AS9090
2	308	14.3	384	2	JC7574
3	308	14.3	385	2	JC7575
4	306	14.2	387	2	B38302
5	303	14.1	388	1	JC7573
6	302.5	14.0	388	1	S19682
7	301	14.0	382	1	PECH
8	299.5	13.9	396	2	A34401
9	298.5	13.8	383	2	A14443
10	298.5	13.8	384	2	A39314
11	296	13.7	387	2	C38302
12	296	13.7	391	2	A43356
13	295.5	13.7	412	1	KHHUD
14	295	13.7	387	2	D38302
15	289.5	13.4	388	1	S19684
16	287.5	13.3	444	2	T24204
17	287	13.3	407	1	KHRTD
18	286	13.3	387	2	E38302
19	285.5	13.2	398	2	S66465
20	285	13.2	398	2	I51185
21	283.5	13.1	388	1	PEMOAR
22	282.5	13.1	388	1	PEHU
23	281.5	13.1	381	1	CMSHR
24	281.5	13.1	388	1	PEMOAJ
25	281.5	13.1	388	2	A30142
26	279.5	13.0	388	2	B30142
27	279.5	13.0	410	1	KHMSD
28	278.5	12.9	386	1	PEPG
29	278	12.9	387	2	JC7245

30	277.5	12.9	380	2	I47176	chymosin (EC 3.4.2
31	277.5	12.9	396	2	S36865	cathepsin E (EC 3.
32	276	12.8	389	2	JE0371	pepsin C (EC 3.4.2
33	273.5	12.7	381	1	CMBO	chymosin (EC 3.4.2
34	270.5	12.5	377	1	PEMOCJ	gastricsin (EC 3.4
35	270.5	12.5	389	2	A38302	pepsin (EC 3.4.23.
36	270	12.5	376	2	I45856	aspartic proteinase
37	268.5	12.5	344	1	KHPGD	cathepsin D (EC 3.
38	267.5	12.4	381	2	JC7247	prochymosin - comm
39	266	12.3	380	2	S03433	canidaepsin (EC
40	266	12.3	405	2	A25379	saccharopepsin (EC
41	264	12.2	396	2	T47207	aspartic proteinase
42	263.5	12.2	388	2	JC7246	pepsinogen C - com
43	262.5	12.2	394	2	B43356	gastricsin (EC 3.4
44	261.5	12.1	387	2	A45117	aspartic proteinase
45	261.5	12.1	388	2	A29937	gastricsin (EC 3.4

ALIGNMENTS

RESULT 1

AS9090

aspartic proteinase (EC 3.4.23.-) BACE precursor - human

N:Alternate names: beta-secretase; beta-site APP cleaving enzyme

C:Species: Homo sapiens (man)

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-May-2000

C:Accession: AS9090

R:Vassar, R.; Bennett, B.D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E.A.; Denis, P.; Teplio

M.A.; Biere, A.L.; Curran, E.; Burgess, T.; Louis, J.C.; Collins, F.; Treanor, J.; Ro

A:Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the tran

A:Reference number: AS9090; MUID:20002972

A>Note: Submitted to Genbank, September 1999

A:Accession: AS9090

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-501 <VAS>

A:Cross-references: GB:AF190725; NID:96118538; PIDN:AAF04142.1; PID:96118539

C:Genetics:

A:Gene: BACE

C:Superfamily: beta-secretase

C:Keywords: Alzheimer's disease; aspartic proteinase; brain; glycoprotein; hydrolase;

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-45/Domain: propeptide #status predicted <PRO>

F:46-501/Product: acid proteinase BACE #status predicted <MAT>

F:461-477/Domain: transmembrane #status predicted <TRN>

F:93-289/Active site: Asp #status predicted

F:153,172,223,354/Binding site: carbohydrate (asn) (covalent) #status predicted

F:330-380/Disulfide bonds: #status predicted

Query Match 100.0%; Score 2156; DB 2; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.6e-173;

Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ETDEPEPEGRGSGFYEMVDNLKSGGYVEMTSGSPQTLNIIIVDGSSNFANGAAP	60
DB	46	ETDEPEPEGRGSGFYEMVDNLKSGGYVEMTSGSPQTLNIIIVDGSSNFANGAAP	105
QY	61	HPFLHRYQOLSTYRDLKRGYVVPYTGKGELGTDLVSIHPGPNVTVRANIAITE	120
DB	106	HPFLHRYQOLSTYRDLKRGYVVPYTGKGELGTDLVSIHPGPNVTVRANIAITE	165
QY	121	SDRFINGNMEGILGLAFAEIRPDDSLPEPFDLVKQTHVNFSLDLCAGAPPLNOS	180
DB	166	SDRFINGNMEGILGLAFAEIRPDDSLPEPFDLVKQTHVNFSLDLCAGAPPLNOS	225
QY	181	EYLAIVGSGMTIGIDHSLYTGSMTPTIRREYVEVIVRVEINGODLKMCKEYNDX	240
DB	226	EYLAIVGSGMTIGIDHSLYTGSMTPTIRREYVEVIVRVEINGODLKMCKEYNDX	285
QY	241	SIYDSGTNLRLPKKVFEEAIVASIKRAASTEKFPDGFVLGEOLVCMQAGTTPWNIIPVIS	300

```

|||||
Db 286 SIYDSGTNLRKRVFAVNAVSKIAASTEKPPDGFGLGEOLVCMQAGTTPNNIPVLS 345
QY 301 LYLMGEVTNOSFRITILPQOYLRPVEDVATSODCYKFAISOSTGTWGAIVMEGFYV 360
Db 346 LYLMGEVTNOSFRITILPQOYLRPVEDVATSODCYKFAISOSTGTWGAIVMEGFYV 405
QY 361 FDRARRKRGFAVSAHVHDEFRTAAVGPVTLDMEDCGINIPQTD 407
Db 406 FDRARRKRGFAVSAHVHDEFRTAAVGPVTLDMEDCGINIPQTD 452

RESULT 2
JC7574
pepsinogen A - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C:Accession: J07574; PC7119
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens
A:Reference number: J07573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: J07574
A:Molecule type: mRNA
A:Residues: 1-384 <IKD>
A:Cross-references: DDBJ:AB045380
A:Accession: PC7119
A:Molecule type: protein
A:Residues: 16-35;57-76 <IK2>
A:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like
C:Genetics:
A:gene: pga
C:superfamily: pepsin
C:keywords: stomach; zymogen

Query Match 14.3%; Score 308; DB 2; Length 384;
Best Local Similarity 27.5%; Pred. No. 4,8e-18;
Matches 99; Conservative 60; Mismatches 133; Indels 68; Gaps 16;

QY 30 YVEMTVGSPPTLNLIVDTGSSNFAVGAPHPFL-----HRYQRLSTYRDLKRG 82
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 72 YGTISIGTPQEFYVIFDTGSSNLMWPSYCCSQACSNHNRFPQOSTFOATNTP 128
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 83 YVYTGKMEGLGDLVSIPIHGPNTVRANIAITESDK-FFINGSNMEGILGAYAE 141
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 129 VSIQYTGSMGFLGYDTGV---GNIOISNOMGISESEPSFLYSPDGLGLAFPS 185
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 142 IARDDSLPEPFDLVKQTHVP-NLFSIQCGAGFPINOSEVLASVGSMLIGIDHSLY 200
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 IA--SSQATPFEDNMWSQGLIPQNLFSYILSSDQ-----QTSYLVFGVDNSY 233
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 201 TGSIMYTPIREWYEVIVVEINGDQL-KMDCKEYNDSIVDSGTNLRPKKVE 258
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 234 SGSLMWVLTLETWQITLDVSVINGVIACSSC-----QALVDGTSLMTGSPPT- 286
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 239 AAVKSIRKASTKFPDGFMLGEOLV-CMQAGTTPWNIFPVISLYLMGEVNSQFRITIL 317
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 287 ANIQWVIGASQDSN-----GQYVINCNNISNMPTIVF-----TIN 321
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 318 PQOYLRPVEDVATSODCYK-FAISOSTGT---VMGAVIMEGFYVFDARRKRIGFA 371
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 322 GVQYPLSPSAIVYKQNOQCGSSGFQAMNLPITNSGDLMLIGDVFIRQYFTVFDPRANNYVAIA 381
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
JC7575
pepsinogen A - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C:Accession: J07575
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.

```

```

J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogen
A:Reference number: J07573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: J07575
A:Molecule type: mRNA
A:Residues: 1-385 <IKD>
A:Cross-references: DDBJ:AB045376
A:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like
C:Genetics:
A:gene: pga
C:superfamily: pepsin
C:keywords: stomach; zymogen

Query Match 14.3%; Score 308; DB 2; Length 385;
Best Local Similarity 27.8%; Pred. No. 4,8e-18;
Matches 99; Conservative 64; Mismatches 133; Indels 60; Gaps 15;

```

```

QY 30 YVEMTVGSPPTLNLIVDTGSSNFAVG---AAPHPFLHRYQRLSTYRDLKRGYV 85
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 YGTISIGTPQEFYVIFDTGSSNLMWPSYCCSPACTNHMFNPQOSTFOATNTPVSI 132
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 86 PYTGKMEGLGDLVSIPIHGPNTVRANIAITESDK-FFINGSNMEGILGAYAEIAR 144
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 133 QYGTGSMGFLGYDTGV---GNIOINQIFGLSDESPGFLYSPDGLGLAFPSILA 188
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 145 PDDSLPEPFDLVKQTHVP-NLFSIQCGAGFPINOSEVLASVGSMLIGIDHSLYTS 203
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 189 -SSQATPFEDNMWSQGLIPQDLFSYVLSOG---GSFVLFEGVDPTSTYGN 237
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 204 LMYPIRREWYEVIVVEINGDQLK--DCKEYNDSIVDSGTNLRPKVEAAV 261
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 238 LNWVLTLETWQITLDVSVISIGQVIACSGC-----SAIVDTGSLMGFP---STPI 287
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 262 KSIRKASTKFPDGFMLGEOLVCMQAGTTPWNIFPVISLYLMGEVTVNSQFRITILPQY 321
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 288 ANIQYVIGANDSNQGV---INCNNISNMPTIVF-----TINQVQ 326
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 322 LRPVED-VATSODDC---YKFAISOSTGT---VMGAVIMEGFYVFDARRKRIGFA 371
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 327 PLPASAVYKQNOQCGSSGFQAMNLPITNSGDLMLIGDVFIRQYFTVFDPRANNYVAMA 382
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
B38302
pepsin (PC 3.4.23.-) II-1 precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Jun-1991 #sequence_revision 20-Sep-1991 #text_change 23-Feb-1997
C:Accession: B38302
R:Kageyama, T.; Tanabe, K.; Kolwai, O.
J. Biol. Chem. 265, 17031-17038, 1990
A:Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nu
A:Reference number: A38302; MUID:91009127
A:Accession: B38302
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-387 <KAG>
A:Cross-references: GB:M59235; GB:J05638
A:Superfamily: pepsin
C:keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion

Query Match 14.2%; Score 306; DB 2; Length 387;
Best Local Similarity 27.1%; Pred. No. 7.2e-18;
Matches 98; Conservative 67; Mismatches 131; Indels 66; Gaps 15;

QY 30 YVEMTVGSPPTLNLIVDTGSSNFAVG---AAPHPFLHRYQRLSTYRDLKRGYV 85
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 75 YGTISIGTPQEFYVIFDTGSSNLMWPSYCCSLACFLKRNPPDSSSTFOATSTSLI 134
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 86 PYTGKMEGLGDLVSIPIHGPNTVRANIAITESD---KFFINGSNMEGILGAYAEI 142
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 135 TYGSGMTGILGYTVKV---GNIEDTNOIGLSKTEPEGITFLV--APFDGILLAVPSI 189
Oy 143 ARPDSLEPFEDSLVKQTHV--PNLFSLOLCAGGAPLNGSEVLASVGSMITIGDHSLYT 201
Db 190 SASDAT--PVPDMMNNEGLVEDLFSVYLSNG-----EKSGVMWGGIDSSYTT 237
Oy 202 GSLMTPTLRREMYEVIIIVRVEINGODIKM--DKEVYDKSIYDSGTTNLRPKKVEEA 259
Db 238 GSLMNPVSHGEGYWOITMDSITINGETIACADSC-----QAVVDTGSLAGPTSAISK 291
Oy 260 AVKSKAASSTEFKPPDGFMLGEOLV--CMQAGTTPMNIFFPVLSYLMGEVYNQSPRITILP 318
Db 292 IQSYIGASKNL-----LGEITISCAIDSLPDIIV-----TINN 325
Oy 319 QOYLRPEVD--VATSODDC---YKFAISQSTGT--VMGAVIMEGFYVFPDRARRIGFAV 372
Db 326 VQYPLPASAVYILKEDDCLSGFDGMNIDTYSGLMILIGDVFIRQYTFVFRANNQVGLAA 385
Oy 373 SA 374
Db 386 AA 387

```

```

RESULT 5
JC7573
pepsinogen C - African clawed frog
N:Alternate names: progastricsin
C:Species: Xenopus laevis (African clawed frog)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasa, S.
J. Biochem. 129, 147-153, 2001
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens
A:Reference number: JC7573; MUID:21064922; PMID:11134969
A:Contents: Stomach
A:Accession: J07573
A:Molecule type: mRNA
A:Residues: 1-383 <IKU>
A:Cross-references: DDBJ:AB045379
A:Accession: PC7118
A:Molecule type: protein
A:Residues: 17-68 <IK2>
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like a
C:enetics:
A:Gene: Pgc
C:Superfamily: pepsin
C:Keywords: stomach; zymogen

```

```

Query Match 14.1%; Score 303; DB 2; Length 383;
Best Local Similarity 27.4%; Pred. No. 1,3e-17;
Matches 104; Conservative 57; Mismatches 122; Indels 96; Gaps 17;

Oy 30 YVEMTVGSPPTLNIIVDTGSSNFAVGA-----APHPLHRYOROLSTYRDLRK 81
Db 67 YVGEISIGPPNPNFLVLEDTGSSNLMVASTYCSQACNTNHP-----FNPSSQSTYSSNQ 122
Oy 82 GYVYVDTQGWGELGTDIVSIPIGPNVTYRANINAIIESDKFLFNGSN-----WEGTLG 136
Db 123 QNSLQGTGSLGILGIDYTVI---QNAVISQGEFGLSETEP---GTNFFYAFAGFDGILG 175
Oy 137 LAYAEIARPDSDLPEFDSLVKQTHV--PNLFSLOLCAGGAPLNGSEVLASVGSMITIGD 196
Db 176 LAYPSIA--VGGATTVMGMMQO---NLLNQP I--FEGYLSGGS--SQNGEVAFAFGVD 225
Oy 197 HSLYGSLSMTPIRREMYEVIIIVRVEINGOD---LKMDCKEYNIDKSIYDSGTTNLRP 253
Db 226 QNYIYGQITMFPVSEIYQWIGIGFSTNGQATGMCSCGC-----QAIYDTGTSLLTAP 279
Oy 254 KKEVFAVKSITKASSTEFKPPDGFMLGEOLVCMQAGTTPMNI--FPVLSYLMG----- 305
Db 280 QSVFSLISQIGAOQDN-----GQYVVSQS-----NIQNLPTISFISVSFPLP 325

```

```

Oy 306 ---EVTNQS-----FRITLPEQYLRPEVDVATSDDCYKFAISQSTGTVMGAVIME 355
Db 326 PSAYVLQSSSGYCTIGMPTLPSQNGPL-----WILGVFLR 364
Oy 356 GFYVFPDRARRIGFAVSA 374
Db 365 EYISYVIDLGNNOVGFATAA 383

```

```

RESULT 6
pepsin A (EC 3.4.23.1) 4 precursor - Japanese macaque
N:Alternate names: pepsinogen A isozyme 4
C:Species: Macaca fuscata (Japanese macaque)
C>Date: 22-Nov-1993 #sequence_revision 19-Oct-1995 #text_change 18-Jun-1999
R:Kageyama, T.; Tanabe, K.; Koiwai, O.
Eur. J. Biochem. 202, 205-215, 1991
A:Title: Development-dependent expression of isozymogens of monkey pepsinogens and st
A:Reference number: S19681; MUID:92037645
A:Accession: S19682
A:Molecule type: mRNA
A:Residues: 1-388 <KAG>
A:Cross-references: EMBL:X59753; NID:938070; PIDN:CAA42425.1; PID:938071
C:Comment: This is a minor component of pepsin at all post-partum stages.
C:Superfamily: Pepsin
C:Keywords: aspartic proteinase; gastric juice; hydrolyase; phosphoprotein; protein di
F:16-388/Product: pepsinogen A 4 #status experimental <ICG>
F:16-62/Domian: activation peptide #status experimental <AP7>
F:63-388/Product: pepsin A 4 #status experimental <EN2>
F:38-39/Cleavage site: Leu-Ile (pepsin) #status experimental
F:62-63/Cleavage site: Leu-Ile (pepsin) #status experimental
F:94,277/Active site: Asp #status predicted
F:107-112,268-272,311-344/Disulfide bonds: #status predicted
F:130/Binding site: phosphate (Ser) (covalent) #status predicted

```

```

Query Match 14.0%; Score 302.5; DB 1; Length 388;
Best Local Similarity 27.6%; Pred. No. 1.4e-17;
Matches 107; Conservative 64; Mismatches 133; Indels 83; Gaps 17;

Oy 3 DEEPEPGRGSPFEMDNLRKSGQGYVEMTVGSPPTLNIIVDTGSSNFAVGAAPHP 62
Db 64 DEQPLE-----NYLDV-----EYFGTIGTRPAQNFYVFDGSSNLWV---PSV 105
Oy 63 FL-----HRYOROLSTYRDLRKGVYVYQGWEGELGTDIVSIPIGPNVTYRANI 115
Db 106 YCYSLACMDHNLFPNPDSSYTRATSKVTSITYGCSMTGILGIDYKV---GGISDTNQT 162
Oy 116 AATRESK--FFINSNNEGILGAYAEIARPDSDLPEFDSLVKQTHV--PNLFSLOLCGA 173
Db 163 FGLSETPGFFELFAPDGLGLAYPSIS--SSGATPVFDMNWRVLSODLFSVYLSAD 220
Oy 174 GFPLNQEVLASVGSWMIIGIDHSLYTGSMTPTLRREMYEVIIIVRVEINGODL--KM 231
Db 221 ----DQS-----GSVYIFGIDSSYTTGSLMNPVSVBSGYWISVDSITMNGKITACAR 270
Oy 232 DCKEYNTDKSIYDSGTTNLRPKKVEEAAYKSITKASSTEFKPPDGFMLGEOLV--CMQAGT 290
Db 271 GC-----QAIYDTGTSLLGTPSPIANIQSDIGASNSD-----GEMVSCAITS 316
Oy 291 TPWNIFPVLSYLMGEVYNQSPRITILPQOY--LRPEVDVATSDDCYK-----FAISQS 344
Db 317 LPLDIV-----TINGVQYPLPASAYILQSGCTSGFQGMVPTESG 358
Oy 345 TGFVMAVIMEGFYVFPDRARRIGFA 371
Db 359 ELWILGDVFIROYTFVFRANNQVGLA 385

```


Db 223 PE--GGAGSELIFGVDHSHFSGLNMVPTKQAYMQIALDNIQVG--TVMFCSE--G 275
 QY 239 DKSIVDSGTTNLR.LPKKVEAAVRSIKRASSTEKPPGCFWLGEDLVCMQAGTTWNIFFV 298
 Db 276 CQALVDITSLITRPSDKIKQLQNAIGAP-----VGEVAVE-----CANLNMVDP 322
 QY 299 ISLILMGVTTNQSFRITLIPQOYLRPVEDVATSODDCYKFAISQSTG----- 346
 Db 323 VTFTING-----VPYTLSPRAY--TLLDPFVDMQFC-----SSGFGGLDHPAPG 366
 QY 347 -TWGAVIMEGEVYVPPARRKRIFGA 371
 Db 367 LMITGDFVIFROFISVFDGRNNRVGLA 392

RESULT 9
 A41443
 pepsin (EC 3.4.23.-) precursor, embryonic - chicken
 C:Species: Gallus gallus (chicken)
 C>Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 21-Jul-2000
 C:Accession: A41443
 R:Hayashi, K.; Agata, K.; Mochi, M.; Yasugi, S.; Eguchi, G.; Mizuno, T.
 J. Biochem. 103, 290-296, 1988
 A:Title: Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken pepsin
 A:Reference number: A41443; M01D:88227903
 A:Accession: A41443
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-383 <HAY>
 A:Cross-references: GB:D00215; NID:q2760810; PIDN:BA00153.1; PID:q222853
 C:Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match 13.8%; Score 298.5; DB 2; Length 383;
 Best Local Similarity 25.2%; Pred. No. 3e-17;
 Matches 90; Conservative 75; Mismatches 125; Indels 67; Gaps 14;

QY 30 YYEMVVGSPQTLNLTIVDTGSSNFAVGA---APHPFLHRYRQRLSTYRDLRKGVYV 85
 Db 76 YGTITIGTPPOFTVFPFGTSSNLMVPSVCSHPACQSHQMFNPQSSTYSGTGNLSI 135
 QY 86 PYTGKMEGELGTDLVSIPIHPGNVTVRANIAITTESDKFFINGSNNEGILGLAYAEIARP 145
 Db 136 HYGTGMEGVGCDIYTVASLMDTNLFGLST-SEPGQFVY-VKFDGILGLGYSLSLA- 192
 QY 146 DDLSEFPDLSLVKQTHV-PNLSLQLCGAGFPLNQSEVLASVGSMTIGIDHSIYTGSL 204
 Db 193 -DGITPFDNMVNESLLEQNLFSVLS-----REPMGSMVFGIDESTFTGSI 240
 QY 205 WTPPIREKRYEVIIVRVEINQDL--KMDCKEYNDKSIYDSGTTNLR.LPKKVEAAVK 262
 Db 241 NMLPVSYGQYQWQISMDSTIYNKQELIACSSGC-----QALIDTGLSLVAGPASIINDIQS 294
 QY 263 SIKAASTTEKFPDGFVWLGELVCMQAGTTPWNIFFVIS---YLMGEVTTNSFRITLIP 318
 Db 295 AVGANQNT-----YGEYSV-----NCSHILAMPVYVFIYIGI----- 326
 QY 319 QOYLARVEVVA---TSQDDCYKFAISQSTGTGTVGAVIMEGEVYVPPARRKRIFGA 371
 Db 327 -OY--PVPALAVTEONGQTCMSFSQNSADLMIIGDVFIRVYSIFDRANNRVGLA 380

RESULT 10
 A39314
 gastricsin (EC 3.4.23.3) precursor - bullfrog
 C:Species: Rana catesbeiana (bullfrog)
 C>Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 22-Jun-1999
 C:Accession: A39314
 R:Yakabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Miki, K.; Kurokawa, K.; Ito, H.; Kageyama, J. Biol. Chem. 266, 22436-22443, 1991
 A:Title: Purification, characterization, and amino acid sequences of pepsinogens and pepsin

A:Reference number: A39314; M01D:92042186
 A:Accession: A39314
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-384 <YAK>
 A:Cross-references: GB:M73750; NID:q213687; PIDN:AAA49530.1; PID:q213688
 C:Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match 13.8%; Score 298.5; DB 2; Length 384;
 Best Local Similarity 25.5%; Pred. No. 3e-17;
 Matches 100; Conservative 61; Mismatches 142; Indels 89; Gaps 15;

QY 14 SFYEMVDNLKSGSGQYVEMTVGSPQTLNLTIVDTGSSNFAVGAHPFL-----HR 66
 Db 51 NFATAFEPLANYMDMSYGEISIGFPQNFVLFPDGSSNLMV---PSYQSQACTNHP 107
 QY 67 YYQRLSTNRYDLRKGVVYPTQKMEGELGTDLVSIPIHPGNVTVRANIA-----AI 118
 Db 108 QRPSSQSSSYSSNQOQFSIQYTGSLTGILGYDVOI-----QNALISQGEFGLSV 158
 QY 119 TESDKFFINGSNNEGILGLAYAEIARPDLSLPIKQTHVFN-LFSQLCGAGFPL 177
 Db 159 TEPGTFVY-AQFDGILGLAYPSIA--EGATTVMQGMILQNLINQPLFAFLYSQOQNSQ 215
 QY 178 NOSEVLASVGSMTIGIDHSIYTGSLWTPPIREMYEVIIVRVEINQD---LKMCK 234
 Db 216 N-----GGEVAVGAGVDQONYSQIYWTPTSETYMQIGIQRSVNGQATGMSQGC- 266
 QY 235 EYNDKSIYDSGTTNLR.LPKKVEAAVRSIKRASSTEKPPGCFWLGEDLV-CWQAGTTPW 293
 Db 267 -----QGIYDITSLITLRPOSVFSLSLMQSIGAQDON-----GQYAVSCSNQSLPT 313
 QY 294 NIFPVI-----SLYLMGEVTTNQ--SFRITLIPQOYLRPVEDVATSODDCYKFAISQ 342
 Db 314 ISFTISGVSPFLPPSAVYLQNSGYCTIGIMPTYLPQNGQPL----- 356
 QY 343 SSTGTVMGAVIMEGEVYVPPARRKRIFGA 374
 Db 357 ----WILDGDFLRQIYVYDIGNNOVGFAA 384

RESULT 11
 C38302
 pepsin (EC 3.4.23.-) II-2/3 precursor - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C>Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 23-Feb-1997
 C:Accession: C38302
 R:Kageyama, T.; Tanabe, K.; Koiwai, O.
 J. Biol. Chem. 265, 17031-17038, 1990
 A:Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nucleotide sequence, and amino acid sequence of the mature enzyme
 A:Reference number: A38302; M01D:91009127
 A:Accession: C38302
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-387 <KAG>
 A:Cross-references: GB:J05638
 C:Superfamily: pepsin
 C:Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion

Query Match 13.7%; Score 296; DB 2; Length 387;
 Best Local Similarity 26.9%; Pred. No. 5e-17;
 Matches 97; Conservative 63; Mismatches 135; Indels 66; Gaps 13;

QY 30 YYVMVVGSPQTLNLTIVDTGSSNFAVGAHPF-----LHRYRQRLSTYRDLRK 82
 Db 75 YFGTISIGTPPOFTVFPFGTSSNLMV---PSYQSSLACALHKRNPEDSSYTGSTSET 131
 QY 83 VYVYPTQKMEGELGTDLVSIPIHPGNVTVRANIAITTESKFFINGSNNEGILGLAYAEI 142
 Db 132 LSLITYGSMTGILGIDTVKVGSIEDTNIQIFGLSKTEPSLTLFL--APFDGILGLAYPSI 189

```

QY 143 ARPDSDLEPFDSLVKQTHV-PNLFSLQICGAGFPLNOSSEVLASVGGSMITGIDHSLYT 201
DB 190 SSSDST--PVIDNMNMNEELVSQDLFSVYLLSSD-----EKGSLVHMGIDSSSYT 237
OY 202 GSLWYPIRREMYEYIVIRVEINGQDILK--DKEYNYDKSIYDSGTTNRLPKKFEFA 259
DB 238 GSLMNVFVSYEGYMOITMDSVINSINETIACADSC-----QAIYDTGTSLLTGP---TS 287
OY 260 AVKSKIAASSTREKFPDGGWLGELQV-CWQAGTTPWNIPFVSLVLMGEVFNQSPRITILP 318
DB 288 AISNIQSYIGASK-----NILEGNVNSGALDLPDIYV-----TING 325
OY 319 OQYLRPEVDVATSDODCYKFAISQSTGT-----VMGAVIMEGFYVVDARRKRGFAV 372
DB 326 IOYLPFASVYILKEDDCTSGLEGNNVDTYTGELMILGDVFIHQYFVFEFRANNQGLIAA 385
OY 373 S 373
DB 386 A 386

```

RESULT 12

```

A43356
A:Title: cathepsin E (EC 3.4.23.34) precursor - guinea pig
N:Alternate names: erythrocyte membrane aspartic proteinase; slow-moving proteinase
C:Species: Cavia porcellus (guinea pig)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: A43356
R:Kageyama, T.; Ichinose, M.; Tsukada, S.; Miki, K.; Kurokawa, K.; Koizumi, O.; Tanji, M.
J. Biol. Chem. 267, 16450-16459, 1992
A:Title: Gastric procathepsin E and procathepsin from guinea pig. Purification, molecu
A:Reference number: A43356; MUID:92355614
A:Molecule type: mRNA
A:Residues: 1-391 <KAG>
A:Cross-references: GB:S52557; NID:9191294; PIDN:AAA37052.1; PID:9191295
A:Note: sequence extracted from NCBI backbone (NCBIN:110763, NCBIPI:110769)
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; membrane protein

```

```

Query Match 13.7% Score 296; DB 2; Length 391;
Best Local Similarity 26.9% Pred. No. 5, 1e-17;
Matches 98; Conservative 64; Mismatches 130; Indels 72; Gaps 16;

```

```

OY 30 YVYEMTVGSPQTLNLTIVDTGSSNFAVGA---APHPFLHRYTQROLSSTYRLKGVYV 85
DB 74 YFGTISIGSPQNETVIFDTGSSNLWVPSYCTSPACQTHPVHPSLSSTYREVGNSTFI 133
OY 86 PYTQGRKEGELGIDLVSIPIRGPNVTVANIAITSDKPFINSNMGGIIGLAVAEIARP 145
DB 134 QYGGSLTGTIGTIGADQVSV-BGLTVVGGQFESVQVEPKTVH-AEPGIIIGLGTPLSAA- 190
OY 146 DDLSEPFDSLVKQTHV-PNLFSLQICGAGFPLNOSSEVLASVGGSMITGIDHSLYTGS 205
DB 191 -GGTVPEFDNMQ-----NLVALPM-----PSYVMSNPGSGSLRFGYDPSHFSGLN 241
OY 206 YTPIRREMYEYIVIRVEINGQDILKMKCKEYNDKSIYDSGTTNRLPKKFEFAVAKRK 265
DB 242 WVPYTKQAYMOIALDGIQV--DSVMFCSE--GCQAIYDTGTSLLTGP-----PEKIQDQ 293
OY 266 AASTETKFPDGGWLGELQV-CWQAGTTPWNIPFVSLVLMGEVFNQSPRITILPQ 320
DB 294 EALGATVYDESY-----SVQC-----ANLNMMLDVT-----FLINGVPTLAMPFA 333
OY 321 YLRPEVDVATSDODCYKFAISQSTGT-----TMGAVIMEGFYVVDARRKRR 367
DB 334 Y--TLDPVDVDMQV-----STGFEGLIOPAPGLWILGDVFIHQYFVFEFRANNR 383
OY 368 IGFA 371
DB 384 VGLA 387

```

RESULT 13

```

KHHUD
A:Title: cathepsin D (EC 3.4.23.5) precursor [validated] - human
N:Alternate names: preprocathepsin D
C:Species: Homo sapiens (man)
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 15-Sep-2000
C:Accession: A25771; S30749; PC2066; I59236; I57716
R:Faust, P.L.; Kornfeld, S.; Chirgwin, J.M.
Proc. Natl. Acad. Sci. U.S.A. 82, 4910-4914, 1985
A:Title: Cloning and sequence analysis of cDNA for human cathepsin D.
A:Reference number: A25771; MUID:85270436
A:Accession: A25771
A:Molecule type: mRNA
A:Residues: 1-412 <FAU>
A:Cross-references: EMBL:M11233; NID:9181179; PIDN:AMB59529.1; PID:9181180
R:Westley, B.R.; May, F.E.B.
Nucleic Acids Res. 15, 3773-3786, 1987
A:Title: Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive human br
A:Reference number: S30749; MUID:87231068
A:Accession: S30749
A:Molecule type: mRNA
A:Residues: 1-412 <WES>
A:Cross-references: EMBL:X05344; NID:929677; PIDN:CA28955.1; PID:929678
R:May, F.E.B.; Smith, D.J.; Westley, B.R.
Gene 134, 277-282, 1993
A:Title: The human cathepsin D-encoding gene is transcribed from an estrogen-regulate
A:Reference number: PC2066; MUID:94085791
A:Accession: PC2066
A:Molecule type: DNA
A:Residues: 1-23 <MA>
A:Cross-references: GB:LI2980; NID:9291930; PIDN:AAA16314.1; PID:9455429
A:Experimental source: MCF-7 cell
R:Cavailles, V.; Augereau, P.; Rochefort, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 203-207, 1993
A:Title: Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate
A:Reference number: I59236; MUID:93126342
A:Accession: I59236
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <CAV>
A:Cross-references: GB:S52557; NID:9263124; PIDN:AAD13868.1; PID:94261856
R:Augereau, P.; Miralles, F.; Cavailles, V.; Gaudeliet, C.; Parker, M.; Rochefort, H.
Mol. Endocrinol. 8, 693-703, 1994
A:Title: Characterization of the proximal estrogen-responsive element of human cathep
A:Reference number: I57716; MUID:95021301
A:Accession: I57716
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <CAV2>
A:Cross-references: GB:S74689; NID:9786350; PIDN:AAD14156.1; PID:94261856
R:Balwin, E.T.; Bhat, T.N.; Guinik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A:Reference number: A51839; PDB:1LYA
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 65-161, 170-241
R:Balwin, E.T.; Bhat, T.N.; Guinik, S.; Hosur, M.V.; Sowder II, R.C.; Cachau, R.E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 6796-6800, 1993
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, with inhibitor residues
A:Reference number: A46229; MUID:93342076
A:Contents: annotation; X-ray crystallography, 2.5 angstroms
C:Comment: Cathepsin D is a ubiquitous lysosomal proteinase.
C:Comment: In addition to the propeptide, residues 163-168 and 411-412 are proteolyti
C:Comment: The carbohydrate bound to 134-Asn contains a mannose-6-phosphate that is b
A:Gene: CTSD
A:Cross-references: GDB:120512; OMIM:116840
A:Map position: 11p15.5-11p15.5
C:Function:

```


QY 177 LNOSEVLASVGSMTIGGIDHSIYTGSLWYTPIRREMYEVIIVREINGODLKMDCREY 236
Db 221 -DOS-----GSVYIFGGIDSSYTGSLMWVPVSEGYWQISVDSITMNGEAI--CAE- 270
QY 237 NYDKSIYDSGFTNRLPKKVEFAVAKSIRAASTEKFPDGFWLGEOLY-CWQAGTTPWNI 295
Db 271 -GGOAIYDTGTSLSLTGFTSPIANIOSDIDGASENSD-----GEMVYSCSAISLDPDIY 321
QY 296 FPVYSILYLMGEVFNOSFRITILPOOYLRPVEDVATISQDDCYK-----FAISQSSTGTVMG 350
Db 322 F-----TINGIQYPPPSAY-----ILQSQSCISGFGMDVPTESGELMILG 364
QY 351 AVIMEGFYVWFDRAKRRIQFA 371
Db 365 DVFTROYFTVFPDRANNOVGILA 385

Search completed: August 7, 2002, 09:14:18
Job time: 42 sec

33	2/0.3	12.3	3/1	1	PEPC_MACFU	P03955	macaca	fusc
----	-------	------	-----	---	------------	--------	--------	------

Chapman C., Gloger I.S., Murphy K.E., Southan C.P., Ryan D.M.,

Smith T.S.,	Simmons D.L.,	Walsh F.S.,	Dingwall C.,	Christie G.
RA				

RT "Identification of a novel aspartic proteinase (Asp 2) as beta-secretase.";
 RL Mol. Cell. Neurosci. 14:419-427(1999).
 RN [5].
 RP SEQUENCE FROM N.A. (ISOFORM B).
 RC TISSUE=Brain, and Pancreas;
 RA Michel B., De Pietri Tonelli D., Zaccchetti D., Keller P.;
 RT "New beta-site APP cleaving enzyme isoform (BACE-1B) obtained from human brain and pancreas";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [6].
 RP SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.
 RX MEDLINE=2014060; PubMed=10677483;
 RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
 CC -1- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP, LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; A/BACE-1A (shown here) and B/BACE-1B; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE EURKARYOTIC ASPARTYL PROTEASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF190725; AAF04142.1; -
 DR EMBL: AF201468; AAF18982.1; -
 DR EMBL: AF200343; AAF17079.1; -
 DR EMBL: AF204943; AAF26367.1; -
 DR EMBL: AF338816; AAK38374.1; -
 DR EMBL: AF200193; AAF13715.1; -
 DR HSSP: P32329; 1YPS.
 DR MIM: 604252; -
 DR InterPro: IPR001969; Asp.-protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp; 3.
 DR PRINTS: PR00792; PEPSIN.
 DR PROSITE: PS00141; ASP-PROTEASE; 1.
 KW Hydrolyase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
 FT Signal; Alternative splicing.
 FT SIGNAL 1 21
 FT PROPEP 22 45
 FT CHAIN 46 501
 FT DOMAIN 22 457
 FT TRANSMEM 458 478
 FT DOMAIN 479 501
 FT ACT_SITE 93 93
 FT ACT_SITE 289 289
 FT CARBOHYD 153 153
 FT CARBOHYD 172 172
 FT CARBOHYD 223 223
 FT CARBOHYD 354 354
 FT VARSPLIC 190 214
 SQ SEQUENCE 501 AA; 55763 MW; 377CE4C82ACEF05 CRC64;
 Query Match 100.0%; Score 2156; DB 1; Length 501;
 Best Local Similarity 100.0%; Pred. No. 1,3e-171;
 Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETDEEPEERGRSGFVEMVDNLRGSGGGYYEMTVGSPPTLNILVDTGSSNFANGAP 60
 DB 46 ETDEEPEERGRSGFVEMVDNLRGSGGGYYEMTVGSPPTLNILVDTGSSNFANGAP 105
 QY 61 HPLHRYQQLSTYRDLAKGYVVPYTGKMGELGTDLVSIPHGPNYVRANIAITE 120
 DB 106 HPLHRYQQLSTYRDLAKGYVVPYTGKMGELGTDLVSIPHGPNYVRANIAITE 165
 QY 121 SDRFFINGSMWEGILGLAYAEIARPPDLSLPPFDSLKQTHVNLFSLOLCGAGFLNOS 180
 DB 166 SDRFFINGSMWEGILGLAYAEIARPPDLSLPPFDSLKQTHVNLFSLOLCGAGFLNOS 225
 QY 181 EVLASVGSMTIGGIDHSLYTGSMTPIREMYEYIYRVELNGODLMDCKEKNYDK 240
 DB 226 EVLASVGSMTIGGIDHSLYTGSMTPIREMYEYIYRVELNGODLMDCKEKNYDK 285
 QY 241 SIYVSGTTLRLPKKVEAAVKSIRKASTKEKPFDFGFWLGEOLVCWQAGTTPMNIPEVIS 300
 DB 286 SIYVSGTTLRLPKKVEAAVKSIRKASTKEKPFDFGFWLGEOLVCWQAGTTPMNIPEVIS 345
 QY 301 LYNGEVTNOSFRITLLPQOYLREVEDVATISODCYFAISOSSTGYMAGVIMEGYVY 360
 DB 346 LYNGEVTNOSFRITLLPQOYLREVEDVATISODCYFAISOSSTGYMAGVIMEGYVY 405
 QY 361 FDRARKRIGFAVSACHVDEFTAAVSGPVTLMEDCGYNIPTQDE 407
 DB 406 FDRARKRIGFAVSACHVDEFTAAVSGPVTLMEDCGYNIPTQDE 452
 RESULT 2
 BACE_RAT 501 AA.
 ID BACE_RAT STANDARD: PRT: 501 AA.
 AC P56819;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
 DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
 DE protease 2) (Asp 2) (Asp2) (Membrane-associated aspartic protease 2)
 DE (Memapsin-2).
 GN BACE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20002972; PubMed=10531052;
 RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
 RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
 RA Fisher S., Fuller J., Curiat E., Burgess T., Louis J.-C., Collins F.,
 RA Treanor J., Rogers G., Citron M.;
 RA "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.";
 RT Science 286:735-741(1999).
 RL
 CC -1- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP, LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE EURKARYOTIC ASPARTYL PROTEASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----

or send an email to license@lsb-sib.ch.

CC EMBL: AF190727; AAF04143.1; -

DR HSSP: P32329; 1YPS.

DR InterPro: IPR001969; Asp_protease.

DR InterPro: IPR001461; Pepsin.

DR Pfam: PF00026; asp. 3.

DR PRINTS: PR00792; PEPsin.

DR PROSITE: PS00141; ASP_PROTEASE; 1.

KW Hydrolyase; Aspartyl protease; Glycoprotein; zymogen; Transmembrane; signal.

FT SIGNAL 1 21 POTENTIAL.

FT PROPEP 22 45 POTENTIAL.

FT CHAIN 46 501 BETA-SECRETASE.

FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 458 478 POTENTIAL.

FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).

FT ACT_SITE 93 93 BY SIMILARITY.

FT ACT_SITE 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 501 AA; 55806 MW; 2484450CB8B7DE3 CRC64;

SEQUENCE

Query Match Best Local Similarity 98.9%; Score 2132; DB 1; Length 501; Matches 400; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

1 ETDEPEPEGRSGSFVEMVNLNKGSGQGYVMTYVSPQTNTILVDTGSSNFANGAP 60

46 ETDEPEPEGRSGSFVEMVNLNKGSGQGYVMTYVSPQTNTILVDTGSSNFANGAP 105

61 HPLRLRYQQLSTTRDLKRGVYVYTGKMGELGTDLVSIPIHPNTVYANAIAT 120

106 HPLRLRYQQLSTTRDLKRGVYVYTGKMGELGTDLVSIPIHPNTVYANAIAT 165

121 SDRFFNGSMWEGILGLAVAEIARPDLSLEPPDSLVKQTHPNFSLDLCAGPPLNGS 180

166 SDRFFNGSMWEGILGLAVAEIARPDLSLEPPDSLVKQTHPNFSLDLCAGPPLNGS 225

181 EVLASVSGKTIIGGIDSLYTGSLWTPYIRRMVYEVIIIVREINGQDLKMKCKEYNDK 240

226 EVLASVSGKTIIGGIDSLYTGSLWTPYIRRMVYEVIIIVREINGQDLKMKCKEYNDK 285

241 SIVDSSTNLRIPKVFEEAASSTKPEPGFVLGEOVLCWQAGTTPWNIFFVIS 300

286 SIVDSSTNLRIPKVFEEAASSTKPEPGFVLGEOVLCWQAGTTPWNIFFVIS 345

301 LYLGEVNTQNSFRITILPQOYLRPVEDVATSDDCYKFAISOSSTGTWAGVIMGEFYV 360

346 LYLGEVNTQNSFRITILPQOYLRPVEDVATSDDCYKFAISOSSTGTWAGVIMGEFYV 405

361 FPRARRKRGFAVSACHVDEFTAVEGFPYTLDMEDCGYNIPTQDE 407

406 FPRARRKRGFAVSACHVDEFTAVEGFPYTLDMEDCGYNIPTQDE 452

RESULT 3

BACE_MOUSE STANDARD; PRT; 501 AA.

AC P56818;

DT 30-MAY-2000 (Rel. 39, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)

DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl

DE (Memapsin-2).

DE (Memapsin-2).

DE BACE.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20002972; PubMed=10531052;

RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,

RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,

RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,

RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,

RA Treanor J., Rogers G., Citron M.,

RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by

RT the transmembrane aspartic protease BACE.";

RL Science 286:735-741(1999).

RN [2]

RP REVISIONS TO 6 AND 81-87.

RA Bennett B.D., Vassar R., Citron M.,

RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=20057170; PubMed=10591213;

RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Torry M.C., Pauley A.M.,

RA Braslier J.R., Strahman N.C., Mathews W.R., Buhl A.E., Carter D.B.,

RA Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;

RT "Membrane-anchored aspartyl protease with Alzheimer's disease

RT beta-secretase activity.";

RL Nature 402:533-537(1999).

CC -1- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE

CC AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF

CC THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,

CC LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED

CC SOLUBLE APP, AND A CORRESPONDING CELL ASSOCIATED CARBOXY-TERMINAL

CC FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE (BY

CC SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: BRAIN.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE

CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>

CC or send an email to license@lsb-sib.ch).

CC -----

DR EMBL: AF190726; AAF04143.2; -

DR EMBL: AF200346; AAF17082.1; -

DR HSSP: P36272; 1AM5.

DR MGD: MGI:1346542; Bace.

DR InterPro: IPR001969; Asp_protease.

DR InterPro: IPR001461; Pepsin.

DR Pfam: PF00026; asp. 1.

DR PRINTS: PR00792; PEPsin.

DR PROSITE: PS00141; ASP_PROTEASE; 1.

KW Hydrolyase; Aspartyl protease; Glycoprotein; zymogen; Transmembrane; signal.

FT SIGNAL 1 21 POTENTIAL.

FT PROPEP 22 45 POTENTIAL.

FT CHAIN 46 501 BETA-SECRETASE.

FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 458 478 POTENTIAL.

FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).

FT ACT_SITE 93 93 BY SIMILARITY.

FT ACT_SITE 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 501 AA; 55747 MW; C085A013145E474E CRC64;

SEQUENCE

Query Match

98.9%; Score 2132; DB 1; Length 501;

Best Local Similarity 98.3%; Pred. No. 1.3e-169;
Matches 400; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETDDEPEPEGRSGFVEMVDNLKSGQGYVEMTVSSPQTINLIYDSSNFAVGAAP 60
D 46 ETDDESEPEGRSGFVEMVDNLKSGQGYVEMTVSSPQTINLIYDSSNFAVGAAP 105
QY 61 HPELHARYQRLSTYRDLRKGVPVPTQCKEDELGTDLVSIIPHGNVTVRANIAATE 120
D 106 HPELHARYQRLSTYRDLRKGVPVPTQCKEDELGTDLVSIIPHGNVTVRANIAATE 165
QY 121 SDKPFINGSMNEGILGLAVAEIARPDSDLEPFDLVKQTHVPLVSIQCGAGFPLNOS 180
D 166 SDKPFINGSMNEGILGLAVAEIARPDSDLEPFDLVKQTHVPLVSIQCGAGFPLNOS 225
QY 181 EYLVASVGGSMIIGTIDHSLYTGSLMYTPIRREMYEVIIVAEINOGDLMKDCKEYNDK 240
D 226 EYLVASVGGSMIIGTIDHSLYTGSLMYTPIRREMYEVIIVAEINOGDLMKDCKEYNDK 285
QY 241 STVDGTTNLRPKVFEAAVKSIAASSTKEDPDGFWLGEOLVCMQAGTTPWNIPIVIS 300
D 286 STVDGTTNLRPKVFEAAVKSIAASSTKEDPDGFWLGEOLVCMQAGTTPWNIPIVIS 345
QY 301 LYLMGEVNTQSFRTITLLPOQYLRPEVDATSDDCYKFAISQSSGTVMGAVIMEGYV 360
D 346 LYLMGEVNTQSFRTITLLPOQYLRPEVDATSDDCYKFAISQSSGTVMGAVIMEGYV 405
QY 361 PPRARRIGFAVSACHVDEFTAAVEGPFVLLDMEDCGYNIPQDE 407
D 406 PPRARRIGFAVSACHVDEFTAAVEGPFVLLDMEDCGYNIPQDE 452

RESULT 4
BAE2 HUMAN STANDARD; PRT; 518 AA.

AC Q9Y520; O9UT6;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta secretase 2 precursor (EC 3.4.23.-) (Beta-site Asp-cleaving
enzyme 2) (Aspartyl protease 1) (Asp 1) (Asp1) (Membrane-associated
aspartic protease 1) (Memapsin-1).
DE BACE2 OR Asp21.
GN Homo sapiens (human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=20051170; PubMed=10591213;
RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA Brashler J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA Tomasselli A.G., Parodi L.A., Heinrichson R.L., Gurney M.E.,
RT "Membrane-anchored aspartyl protease with Alzheimer's disease
beta-secretase activity".
RT Nature 402:533-537(1999).
RL [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Bone marrow;
RC Xin H., Stephens J.C., Duan X., Harrowe G., Kim E., Grieshammer U.,
RA Gliese K.;
RT "Identification of a novel aspartic-like protease differentially
expressed in human breast cancer cell lines".
RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RL [3]
RN SEQUENCE FROM N.A.
RP Acciarini M.P., Fumagalli P., Ottolenghi S., Taramelli R.;
RT "Cloning of a gene from chromosome 21 Down region encoding a potential
transmembrane aspartyl protease".
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RL [4]
RN SEQUENCE FROM N.A.
RP Solans A., Estivill X., de la Luna S.;

RT "Cloning of a novel mammalian aspartyl protease."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL [5]
RN SEQUENCE FROM N.A.
RP MEDLINE=20030166; PubMed=10561122;
RX Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
RA Chapman C., Glover I.S., Murphy K.E., Southan C.D., Ryan D.M.,
RA Smith T.S., Simons D.L., Walsh F.S., Dingwall C., Christie G.;
RT "Identification of a novel aspartic proteinase (asp 2) as
beta secretase".
RT Mol. Cell. Neurosci. 14:419-427(1999).
RL [6]
RN SEQUENCE FROM N.A.
RP MEDLINE=2014406; PubMed=10677483;
RX Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of
beta-amyloid precursor protein".
RT Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
RL [7]
RN SEQUENCE FROM N.A.
RP MEDLINE=20289799; PubMed=10830953;
RX Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Menzel U., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Blocker H.,
RA Kanster J., Beck A., Klages S., Hennig S., Riesseltmann L., Dagand E.,
RA Weinmeyer S., Borzym K., Gardiner K., Yaspo M.-L.;
RA Lehnach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21."
RL Nature 405:311-319(2000).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF200342; AAF17078.1;
CC EMBL: AF117892; AAD45240.1;
CC EMBL: AF050171; AAD45963.1;
CC EMBL: AF178532; AAF29494.1;
CC EMBL: AF204944; AAF26368.1;
CC EMBL: AF200192; AAF13714.1;
CC EMBL: AL163284; CAB90458.1;
CC EMBL: AL163285; CAB90554.1;
CC HSSP: P00797; ZREN.
CC MIM: 605668;
CC InterPro: IPR001969; Asp-protease.
CC InterPro: IPR001461; Pepsin.
CC Pfam: PR00026; asp. 3.
CC PRINTS: PR00792; Pepsin.
CC PROSITE: PS00141; ASP_PROTEASE. 2.
CC KEGG: Hydroxylase; Aspartyl protease; Glycoprotein; zymogen; Transmembrane;
CC Signal.
CC FT SIGNAL 1 20
CC FT PROPEP 21 518
CC FT CHAIN ? 518
CC FT DOMAIN 21 473
CC FT DOMAIN 474 494
CC FT DOMAIN 495 518
CC FT ACT_SITE 110 110
CC FT ACT_SITE 303 303
CC FT ACT_SITE 170 170
CC CARBOHYD 170 170
CC -----
CC POTENTIAL.
CC BETA SECRETASE 2.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC BY SIMILARITY.
CC N-LINKED (GLCNAC. .) (POTENTIAL).

RESULT	5			
PEP1_GADMO		STANDARD:	PRT:	324 AA.
AC	PEP1_GADMO			
DT	P56272;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DE	15-JUL-1998 (Rel. 36, Last annotation update)			
OS	Gadus morhua (EC 3.4.23.1).			
OC	Gadus morhua (Atlantic cod).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Paracanthopterygii; Gadiformes; Gadoidae; Gadidae;			
OX	Gadus.			
OX	NCBI_TaxID=6049;			
RN	[1]			
RP	SEQUENCE AND X-RAY CRYSTALLOGRAPHY.			
RC	TISSUE=Stomach;			
RA	Karlsten S., Hough E., Olsen R.L.;			
RT	"Structure and proposed amino-acid sequence of a pepsin from Atlantic			
RT	cod (Gadus morhua)."			
RL	Acta Crystallogr. D 54:32-46(1998).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE			
CC	EUKARYOTIC ASPARTYL PROTEASES FAMILY.			
DR	PDB: 1AMS; 24-DEC-97.			
DR	InterPro: IPR001969; ASP. Protease.			
DR	InterPro: IPR001461; Pepsin.			
DR	Pfam: PF00026; asp. 1.			
DR	PRINTS: PR00792; PEPsin.			
DR	PROSITE: PS00141; ASP. PROTEASE. 2.			
KW	Hydrolase; Aspartyl protease; Digestion; Stomach; Gastric juice;			
KM	3D-structure.			
FT	ACT_SITE	32	BY SIMILARITY.	
FT	ACT_SITE	214	BY SIMILARITY.	
FT	DISULFID	45	BY SIMILARITY.	
FT	DISULFID	206	BY SIMILARITY.	

```

QY 18 MVDNNGKSSQGYEYEMVYGSPPQTLNLTVDTGSSNFAG---AAPHPFLHRYQORLS 73
Db 2 VTQOMKNEADTELEYGSIIGTPPESEFKVIFPTDGTSSNLMTWSSSHCSQACSNNHKEKPPQS 61
QY 74 STYRDLRKGYVPYQYQKKKEGSELGTDLVSIPHG--PWYTRAIATAITSSDKFFLINGSNW 13
Db 62 STYVEKGKTVLDLTGTGKMRGILGDVTVSVGGSGDPQOELG---ESQTEPPGPOA-AAPF 11
QY 132 EGILIGLVAEIAIRAPDSDLEPFEDSLVKQTHY-PULFSLQCGAPFLNSELVASVGSM 19
Db 118 DGLTGLATYSIAA--GAVPVPFNDNGSSLSLEAKDSEFFYLSGGG--ANSSEW----- 16
QY 191 ITGGIDHSLYTGSLWYTPPIREMYEYVLIARVELINGDGLKMD-CKEYNDKSIYDSCGTTN 24
Db 167 -LEGVNSNYHTYSIHMIPVTAETAKYQVVALDGIIVNGTACBEC-----QAVDTGTSK 21
QY 250 LRAPKVEFLAAVKSLKASSSTEEKPPDGFWMJGEOLVCQOAGTTPMNIPEYISLYLGEVYN 30
Db 220 IYAPVASALANIMKDIGASEN-----QOEMMGN--CASVOSLDPITF-----TI 26
QY 310 QSEFRILILPOQLIREPEVQATSDODCYKFAISQSSGT-----YMGAVIMEGEYVF 36
Db 261 NGVKGQPLPSASVIEDQAFCS-----GLGSSVPSNISELMTIFDVFRLNRYTYI 31
QY 362 DRARRKRIGFAVSA 374
Db 312 DRTNNKVGFAFAA 324

```

```

RESULT      6
CATD_BOVIN
AC          ID          CATD_BOVIN          STANDARD;          PRT;          390 AA.
AC          P80209; Q9TS27;
DT          01-JUL-1993 (Rel. 26, Created)
DT          16-OCT-2001 (Rel. 40, Last sequence update)
DT          16-OCT-2001 (Rel. 40, Last annotation update)
DE          Cathepsin D precursor (EC 3.4.23.5).
GN          CTSD.
OS          Bos taurus (Bovine).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OX          Bovidae; Bovinae; Bos.
NX          NCBI_TaxID=9913;
RN          [1]
RP          SEQUENCE OF 1-48.
RC          TISSUE=Milk;
RC          MEDLINE=93202276; PubMed=8454061;
RA          Larsen L.B., Bolsen A., Petersen T.E.;
RT          "Procathepsin D cannot autocatalyze to cathepsin D at acid pH.";
RL          FEBS Lett. 319:54-58(1993).
RN          [2]
RP          SEQUENCE OF 45-390, AND X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RC          TISSUE=Liver;
RC          MEDLINE=93223670; PubMed=8467789;
RA          Metcalf P., Fusek M.;
RT          "Two crystal structures for cathepsin D: the lysosomal targeting
RT          signal and active site.";
RL          EMBO J. 12:1293-1302(1993).
CC          -1- FUNCTION: CATHEPSIN D IS AN ACID PROTEASE ACTIVE IN INTRACELLULAR
CC          PROTEIN BREAKDOWN.
CC          -1- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC          that of pepsin A. Does not cleave the 4-Gln-1-His-5 bond in B
CC          chain of insulin.
CC          -1- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC          -1- SUBCELLULAR LOCATION: Lysosomal.

```

Query Match	Best Local Similarity	Score	DB 1;	Length
Matches 98; Conservative	27.1%;	306;	DB 1;	387;
	67; Mismatches 131; Indels 66; Gaps 15			
30 YVEMTVGSPQTLNLVDGSSNFVNG---AAPHELRHYQRLSSYRDLRKGVY 85	YVEMTVGSPQTLNLVDGSSNFVNG---AAPHELRHYQRLSSYRDLRKGVY 85			
75 YFETISIGPPQEFYVIFDTGSSNLWPSTYCSLACFLHKRNPDDSTFQATSETLSI 134	YFETISIGPPQEFYVIFDTGSSNLWPSTYCSLACFLHKRNPDDSTFQATSETLSI 134			
86 PYQGWKEGLGDLVSIPIGPNVTVRANIMATIESD--KEPINGSNWEGTIGLAYAEI 142	PYQGWKEGLGDLVSIPIGPNVTVRANIMATIESD--KEPINGSNWEGTIGLAYAEI 142			
135 TYVGTGSMTGILGVDYKVV--GNIEDPTNQFGLSKTEPGITFLV--APFDGIIGLAPSI 189	TYVGTGSMTGILGVDYKVV--GNIEDPTNQFGLSKTEPGITFLV--APFDGIIGLAPSI 189			
143 ARPDDSLPEFDSLVKQTHV-PULFSIQLCGAGFPLNQSEVLASVGSMSIIGIDHSLYT 201	ARPDDSLPEFDSLVKQTHV-PULFSIQLCGAGFPLNQSEVLASVGSMSIIGIDHSLYT 201			
190 SASDAT--PVEDMMWMDGLVSEDLFVYLLSNG-----EKGSMVWFGGIDSSYTT 237	SASDAT--PVEDMMWMDGLVSEDLFVYLLSNG-----EKGSMVWFGGIDSSYTT 237			
202 GSIMVPIPIREWEYEVIVVEINGDGLK--DCKEYNYDKSIYDSGSTMRLRPKVVEA 259	GSIMVPIPIREWEYEVIVVEINGDGLK--DCKEYNYDKSIYDSGSTMRLRPKVVEA 259			
238 GSIAMVPIVSEGWQITMDSITNGELIACADSC-----QAVVDITGSLAGPTSAISK 291	GSIAMVPIVSEGWQITMDSITNGELIACADSC-----QAVVDITGSLAGPTSAISK 291			
260 AVKSIKASSTEFKPDGFWLGEOLV-CWQAGTTPWNIPIYISLYLMEVYNQSRITLLP 318	AVKSIKASSTEFKPDGFWLGEOLV-CWQAGTTPWNIPIYISLYLMEVYNQSRITLLP 318			
292 IQSYIGASKNL-----LGENIISCSAIDSLPDIYF-----TINN 325	IQSYIGASKNL-----LGENIISCSAIDSLPDIYF-----TINN 325			
319 QYQIRPEV--VATSDQD---YKRAISQSSNGT--WGAIVMEGFVYVDPARKRIGFAY 372	QYQIRPEV--VATSDQD---YKRAISQSSNGT--WGAIVMEGFVYVDPARKRIGFAY 372			
326 VQYPLPASAVILKEDDCLSGFDGMNDITSYGELMITIGDVFIRQYTFVFDPRANNQVGLAA 385	VQYPLPASAVILKEDDCLSGFDGMNDITSYGELMITIGDVFIRQYTFVFDPRANNQVGLAA 385			
373 SA 374	SA 374			
386 AA 387	AA 387			

```

AC P27678:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Pepsin A-4 precursor (EC 3.4.23.1) (Pepsin I/II).
GN PCA.
OC Macaca fuscata fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9543;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-70.
RC TISSUE=gastric mucosa.
RX MEDLINE=92037645; PubMed=1935977;
RA Kageyama T., Tanabe K., Kolwai O.;
RT "Development-dependent expression of isozymes of monkey
  pepsinogens and structural differences between them."
RL Eur. J. Biochem. 202:205-215(1991).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY, ALTHOUGH BONDS
  INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
  ALSO CLEAVED TO SOME EXTENT.
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE PHE-, LEU-.
CC -1- MISCELLANEOUS: THE EXPRESSION OF PEPINOGEN GENES IS REGULATED BY
  HORMONES AND RELATED SUBSTANCES.
CC -1- MISCELLANEOUS: EACH PEPINOGEN IS CONVERTED TO CORRESPONDING
  PEPsin AT PH 2.0 IN PART AS A RESULT OF THE RELEASE OF A 47 AA
  ACTIVATION SEGMENT AND IN PART AS A RESULT OF STEPWISE PROTEOLYTIC
  CLEAVAGE VIA AN INTERMEDIATE FORM(S).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
  EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to: license@isb-sib.ch).
CC -----
DR EMBL: X59753; CAA42425.1; -.
DR PIR: S16065; S16065.
DR PIR: S19682; S19682.
DR HSSP: P00790; IPSN.
DR MEROPS: A01.001; -.
DR InterPro: IPR00169; Asp_protease.
DR InterPro: IPR001691; Laminin_G.
DR Pfam: PF00026; asp. 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
KW Hydroxylase; Aspartyl protease; Digestion; Stomach; Gastric juice;
KW Zymogen; Multigene family; Signal; Glycoprotein.
FT SIGNAL 1 15
FT PROPEP 16 38 ACTIVATION PEPTIDE.
FT PROPEP 39 62 ACTIVATION PEPTIDE.
FT CHAIN 63 388 PEPsin A-4.
FT ACT_SITE 94 94 BY SIMILARITY.
FT ACT_SITE 277 277 BY SIMILARITY.
FT DISULFID 107 112 BY SIMILARITY.
FT DISULFID 268 272 BY SIMILARITY.
FT DISULFID 311 344 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .).
SQ SEQUENCE 388 AA; 4195 MW; A2923AB1F7FCDDB9 CRC64;

```

```

Query Match 14.0%; Score 302.5; DB 1; Length 388;
Best Local Similarity 27.6%; Pred. No. 1e-17;
Matches 107; Conservative 64; Mismatches 133; Indels 83; Gaps 17;
OY 3 DEEPEERGRSFEVMDNLRKSGGQGYVMTVGSPPQTLNIIIVDTGSSNFAVGAAPHP 62
DB 64 DEQPLE-----NYLDV-----EYFGTIGTGAQNTVFVDGSSNLTAV---PSV 105

```

```

OY 63 FL-----HRYQRLSTYRDLKRGVVPYTGKWESELTDLVSIPIHGNTVRANI 115
DB 106 YCYSIACMDHNFENPDSSSTRATSKTWSITGTGSMGILDTIVK---GGISDNTDI 162
OY 116 AAITESDK-FFINGSNWEIILGLAVAEIARPPDSLPEFPDSLYKOTY-PNIFSQLCGA 173
DB 163 FGLSETEPGLFLKAPFEDGILALAPSS--SSGATPVFDNWNRLVSQDLFSYLSAD 220
OY 174 GFLNDSVYLVASGSMIIGIDHSLYGLSLWYTPIRREYVEYIIVVEINGODL--KM 231
DB 221 ---DQS-----GSVYIFGIDSSYTGSLNWPVSVYGWQISVDITNNGKTIACAK 270
OY 232 DCKEYNYKSIYDSTTNRLPKKVFEEAIVSIRKASSTKFPDGFUIGEDLV-CWQAGT 290
DB 271 GC-----QALVDITSLTGTSPSIANIISDIGASENSD-----GEMVSSCAISS 316
OY 291 TPNWIFPVYSLYIMGEVYNOSFRITLPQY-LRPVEDVATSDQCYK-----FAISQSS 344
DB 317 LFDIVF-----TINGVQYPLPSAIILOSQSGCTSGFGMDVPIESG 358
OY 345 TGTVGAVIMEGFYVDFRARRIGFA 371
DB 359 ELWILGDVFIROYFTVDFRANNOVGLA 385

```

```

RESULT 9
PEPA_CHICK STANDARD; PRT; 367 AA.
AC P00793;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Pepsin A precursor (EC 3.4.23.1).
OS Gallus gallus (chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=84004412; PubMed=6617663;
RA Baudys M., Kostka V.;
RT "Covalent structure of chicken pepsinogen."
RL Eur. J. Biochem. 136:89-99(1983).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY, ALTHOUGH BONDS
  INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
  ALSO CLEAVED TO SOME EXTENT.
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE PHE-, LEU-.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
  EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC PIR: A00984; PECH.
CC HSSP: P00794; ACMS.
CC MEROPS: A01.001; -.
CC InterPro: IPR00169; Asp_protease.
CC InterPro: IPR001461; Pepsin.
CC Pfam: PF00026; asp. 1.
CC PRINTS: PR00792; PEPsin.
CC PROSITE: PS00141; ASP_PROTEASE; 2.
KW Hydroxylase; Aspartyl protease; Digestion; Stomach; Zymogen;
KW Glycoprotein; Gastric juice.
FT PROPEP 1 42 ACTIVATION PEPTIDE.
FT CHAIN 43 367 PEPsin A.
FT ACT_SITE 77 77
FT ACT_SITE 260 260
FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .).
FT DISULFID 90 95
FT DISULFID 251 255
FT DISULFID 290 323
SQ SEQUENCE 367 AA; 40431 MW; 0C547E7FD8F5B341 CRC64;

```

```

Query Match 14.0%; Score 302; DB 1; Length 367;

```

Best Local Similarity 24.0%; Pred. No. 1e-17;
Matches 88; Conservative 69; Mismatches 125; Indels 84; Gaps 13;

```

OY 30 YVEMTVGSPPTQINILVDGSSNFVANGAHPRL-----HRYGROASTFYRDLKRG 82
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 YGRTISIGTPOODSVIFDTGSSNLMW---PSYCKSSACSNHRFPDPSSTIYVSTNET 115
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 83 VVYVYTGKMGEGELGTDVLSIPHGPNVYVRANIAITESDK-FFINGSNMEGILGLAYAE 141
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 116 VYVAVGTSMSGILIGYDVAV---SSIDVQNOIGLSTEGSEFFYCNFGDIGLAFPS 172
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 142 IARPDLSLEPFDSLVKQTHV-PNLFSLQLCGAGFPPLNQSEVLASVGSMTIGGIDISLY 200
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 173 IS--SSGATPVFDNMMSOHLVAODLFFSVYLSKDG-----ETGSEVLEFGIDPYVT 220
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 201 TGSIMYTFIREMYEVIIVREINGODLK--MDCKEYNDKSIYDSGTMLRLPKVFE 258
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 221 TKGIVWPLSAETWQITWDRYVGNKYVACEFTC-----QATVDIGTSLVMDQAGYN 274
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 259 AAVKSIRKASSTE-----KFPDGFMLGEOLVCMQAGTTPWNIFFPVISLYMEVNTQS 311
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 275 RIKNLGVSSDDEISCDISKLPD-----VTFHNGHA--- 307
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 312 FRITTLPOOYLPRVEDVATSDQCKKFAISQSGT-----VMGAVIMEGFYVVDRAR 365
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 308 -----FTLPASAYVLNEDGSCMLGFENMGTFELGEOMILDVFIREFYVIFEDKAN 358
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 366 KRIGFA 371
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 NKVGLS 364
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 10
CATE HUMAN STANDARD; PRT; 396 AA.
ID CATE HUMAN STANDARD; PRT; 396 AA.
AC P14091;
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cathepsin E precursor (EC 3.4.23.34).
GN CTSE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89380302; PubMed=2674141;
RA Azuma T., Pals G., Mohandas T.K., Couvreur J.M., Taggart R.T.;
RT "Human gastric cathepsin E. Predicted sequence, localization to
RT chromosome 1, and sequence homology with other aspartic
RT proteinases."
RT J. Biol. Chem. 264:16748-16753(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112877; PubMed=1370478;
RA Azuma T., Liu W.G., Vander laan D.J., Bowcock A.M., Taggart R.T.;
RT "Human gastric cathepsin E gene. Multiple transcripts result from
RT alternative polyadenylation of the primary transcripts of a single
RT gene locus at 1q31-q32."
RL J. Biol. Chem. 267:1609-1614(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Tanneil P.J., Kay J.;
RT "Human procathepsin E."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 54-68; 77-95; 141-155; 275-285 AND 389-396.
RX MEDLINE=90241267; PubMed=2334440;
RA Athauda S.B.P., Matsuzaki O., Kageyama T., Takahashi K.;
RT "Structural evidence for two isozytic forms and the carbohydrate
RT attachment site of human gastric cathepsin E."
RL Biochem. Biophys. Res. Commun. 168:878-885(1990).

```

```

CC -1- FUNCTION: DUE OF ITS INTRACELLULAR LOCATION AND DISTRIBUTION IN
CC LYMHOID ASSOCIATED TISSUE, IT MAY HAVE A ROLE IN IMMUNE FUNCTION.
CC -1- CATALYTIC ACTIVITY: Similar to cathepsin D, but slightly broader
CC specificity.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL; M84424; AAA52300.1; -
DR EMBL; M84413; AAA52300.1; JOINED.
DR EMBL; M84417; AAA52300.1; JOINED.
DR EMBL; M84418; AAA52300.1; JOINED.
DR EMBL; M84419; AAA52300.1; JOINED.
DR EMBL; M84420; AAA52300.1; JOINED.
DR EMBL; M84421; AAA52300.1; JOINED.
DR EMBL; M84422; AAA52300.1; JOINED.
DR EMBL; J05036; AAA52130.1; -
DR EMBL; A1250717; CAB82850.1; -
DR PIR; A34401; A34401.
DR PIR; A34643; A34643.
DR PIR; A42038; A42038.
DR HSSP; P00794; 4CMS.
DR MEROPS; A01.010; -.
DR MIM; 116890; -.
DR InterPro; IPR001969; Asp_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolyase; Aspartyl protease; Glycoprotein; Zymogen; signal.
FT SIGNAL 1 17
FT PROPEP 18 53
FT CHAIN 54 396
FT MOD_RES 18 18
FT ACT_SITE 96 96
FT ACT_SITE 281 281
FT DISULFID 60 60
FT DISULFID 109 114
FT DISULFID 272 276
FT DISULFID 314 351
FT CARBOHYD 90 90
FT CARBOHYD 220 220
FT CARBOHYD 333 333
SQ SEQUENCE 396 AA; 42793 MW; 40B643C5FB01521E CMC64;

```

Query Match 13.9%; Score 299.5; DB 1; Length 396;
Best Local Similarity 25.9%; Pred. No. 1.9e-17;
Matches 100; Conservative 65; Mismatches 148; Indels 73; Gaps 16;

```

OY 3 DEEPPEPGRGGSFVENDLNRKSGGQGYVEMTVGSPPTQINILVDGSSNFAVA----- 58
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 DQSAKEP-----LINYLD-----MEYFGTISGSPDNFTVHIDTGSNLMWPSVYCT 110
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 59 AHPHFLHRYQRLSTFYDRLKRGVVPYPTQCKMGEELGTDVLSIPHGPNVYVRANIAI 118
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 111 SPACKTHSRFPQSSSTYQOPGQSFISIQYGTSLSSITADQVSV-EGLTIVGQGFGESEV 169
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 119 TESDKFFINGSNMEGILGLAYAEIARPDLSLEPFDSLVKQTHVPLNLFSLQLCGAGFPLN 178
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 TEGGQGFVD-AERDGLIGLCYPSLA--VGGVYVPVFNMAA-----NLVPLPMFSYVMSN 222
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 179 QSEVLASVGSMTIGGIDLSLYTGSIMYPIREMYEVIIVREINGODLKMDCKEYV 238
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 PE--GGAGSELIIFGVDHSHFSGSLNWVPVTKQAVQIALDMDIIVGG--TYVFCSE--G 275
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


QY 239 DKSIVDSSTNLRLPKRYEAAVKSIRKASSTKFFPDGFMLGQVLCWQACTPWNIFPV 298
 Db COATVDSTSLITGSPDKIKLOLNAIGAAP-----VDESEVAVE-----CANLWMPD 322
 QY 299 ISLYLMEGVETNQSRRITLIPQOYLPRVEDVATSDDDCKFAISQSSFG----- 346
 Db 323 VTFING-----VPYTLSPAY--TLIDFVDMQFC-----SSGFGOLDIHPPAGP 366
 QY 347 -TVNGAVIMEGFVYVDFARRRIGFA 371
 Db 367 LMIIGDVFIHQFVSFVDFGRNNRVGLA 392

RESULT 11
 PEPE_CHICK STANDARD: PRT: 383 AA.
 AC P16476;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Embryonic pepsinogen precursor (EC 3.4.23.-)
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88227903; PubMed=3131317;
 RA Hayashi K., Agata K., Mochi M., Yasugi S., Eguchi G., Mizuno T.;
 RT "Molecular cloning and the nucleotide sequence of cDNA for embryonic
 chicken pepsinogen: phylogenetic relationship with prothymosin";
 RL J. Biochem. 103:290-296(1988).
 CC -1- DEVELOPMENTAL STAGE: SPECIFICALLY SECRETED DURING THE EMBRYONIC
 CC PERIOD IN THE CHICKEN PROVENTRICULUS (GLANDULAR STOMACH).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: D00215; BAA00153.1; -.
 DR PIR: A41443; A41443.
 DR HSSP: P00794; 4CMS.
 DR MEROPS: A01.028; -.
 DR InterPro: IPR001969; Asp_protease.
 DR InterPro: IPR001461; Pepsin.
 DR Pfam: PF00026; asp. 1.
 DR PRINTS: PRO0792; PEPsin.
 DR PROSITE: PS00141; ASP_PROTEASE; 2.
 DR KMW HydroLase; Aspartyl protease; Digestion; Signal; Glycoprotein.
 FT SIGNAL 1 16
 FT CHAIN 17 383
 FT ACT_SITE 94 94
 FT ACT_SITE 276 276
 FT DISULFID 107 112
 FT DISULFID 267 271
 FT DISULFID 310 344
 FT CARBOHYD 132 132
 FT CARBOHYD 204 204
 FT CARBOHYD 309 309
 FT CARBOHYD 350 350
 FT VARIANT 51 51
 SO SEQUENCE 383 AA; 41719 MW; 1642796871611F54 CRC64;

Query Match 13.8%; Score 298.5; DB 1; Length 383;

Best Local Similarity 25.2%; Pred. No. 2.2e-17;
 Matches 90; Conservative 75; Mismatches 125; Indels 67; Gaps 14;

QY 30 YVEMTVGSPPTNLINVTGSSNFAVGA-----APRFLHRYXQRLSTYDLKQYV 85
 Db 76 YVGTISIGTPDPOFVYVFDTGSSNLWVPVSCSTPACOSHOMENPSQSTYSTSQONLSI 135
 QY 86 PYTGKMEGELGTDVLSIPHGPVYVRANIAMITESDKFFINGSMWEGILGLAVAEIARP 145
 Db 136 HGTGDMEGTVCDDVTVASLMDTNQLFELST-SEPGQFFV-VAFDGLIGLGYSLA- 192
 QY 146 DDLSEFPDSIVKQTHV-PNLSLQLCGAGFPLNOSVLAISGSMIIIGDLSLYTGS 204
 Db 193 -DGITVFEDNMWNESSLLEONLEFVYLS-----REPMSVYVGGIDSEYFGTSI 240
 QY 205 WYPIRREWYEVYIVRYVINGODL--KMDCKEYVDKSYVSGTNTLRPKRYEAAV 262
 Db 241 NMIPVSYQGYWQISMDSTIVNKOELACSSGC-----QHLIDTGLVAGPSDINDIOS 294
 QY 263 SIKASSTKFPDGFMLGQVLCWQAGTTPWNIFPVISL---YLMGEVTVNQSFRTILP 318
 Db 295 AVGANQNT-----YGEYSV-----NCHILMPDVVFYIGI----- 326
 QY 319 QOYLPRVEDVA---TSDDCKFAISQSSSTGTVMGAVIMEGFVYVDFARRRIGFA 371
 Db 327 -QY--PVRLATTEQNGGCTCMSSFONSSADLMIIGDVFIHQFVSFVDFGRNNRVGLA 380

RESULT 12
 PEP2_RABIT
 ID PEP2_RABIT STANDARD: PRT: 387 AA.
 AC P27821;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Pepsin II-2/3 precursor (EC 3.4.23.1) (Pepsin A).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91009127; PubMed=2129536;
 RA Kageyama T., Tanabe K., Koizumi O.;
 RT "Structure and development of rabbit pepsinogens. Stage-specific
 RT zymogens, nucleotide sequences of cDNAs, molecular evolution, and
 RT gene expression during development.";
 RL J. Biol. Chem. 265:17031-17038(1990).
 CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS
 CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
 CC ALSO CLEAVED TO SOME EXTENT.
 CC -1- CATALYTIC ACTIVITY: PREPERENTIAL CLEAVAGE PHE-, LEU-.
 CC -1- DEVELOPMENTAL STAGE: PEPSINOGENS IN GROUP I, II, AND III WHERE
 CC THE PREDOMINANT ZYMOGENS AT LATE POSTNATAL STAGE.
 CC -1- MISCELLANEOUS: THE EXPRESSION OF PEPSINOGEN GENES IS REGULATED BY
 CC HORMONES AND RELATED SUBSTANCES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
 CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M59225; AAA85369.1; -.
 DR PIR: C38302; C38302.
 DR HSSP: P00790; IPSN.
 DR MEROPS: A01.001; -.
 DR InterPro: IPR001969; Asp_protease.
 DR InterPro: IPR001461; Pepsin.


```

Db      334 Y--TLDEVDGKQVY-----STGEESLEIQPPAGPLMLDVFIRQFVAFVDRGNMR 383
QY      368 IGFA 371
Db      384 VGLA 387

RESULT 14
CATD_HUMAN
ID      CATD_HUMAN      STANDARD:      PRT:      412 AA.
AC      P07339;
DT      01-APR-1988 (rel. 07, Created)
DT      01-APR-1988 (rel. 07, Last sequence update)
DT      01-MAR-2002 (rel. 41, Last annotation update)
DE      Cathepsin D precursor (EC 3.4.23.5).
GN      CTSD.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      (1)
RP      SEQUENCE FROM N.A.
RX      MEDLINE=85270436; PubMed=3927292;
RT      Faust P.L., Kornfeld S., Chirgwin J.M.;
RT      "Cloning and sequence analysis of cDNA for human cathepsin D.";
RL      Proc. Natl. Acad. Sci. U.S.A. 82:4910-4914(1985).
RN      (2)
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87231068; PubMed=3588310;
RT      Westley B.R., May F.E.B.;
RT      "Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive
RT      human breast cancer cells.";
RL      Nucleic Acids Res. 15:3773-3786(1987).
RN      (3)
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91299158; PubMed=2069717;
RT      Redecker B., Heckendorf B., Grosch H.W., Mersmann G., Hasilik A.;
RT      "Molecular organization of the human cathepsin D gene.";
RL      DNA Cell Biol. 10:423-431(1991).
RN      (4)
RP      SEQUENCE OF 1-22 FROM N.A.
RX      MEDLINE=94085791; PubMed=8262386;
RT      May F.E., Smith D.J., Westley B.R.;
RT      "The human cathepsin D-encoding gene is transcribed from an estrogen-
RT      regulated and a constitutive start point.";
RL      Gene 134:277-282(1993).
RN      (5)
RP      SEQUENCE OF 1-22 FROM N.A.
RX      MEDLINE=95021301; PubMed=7935485;
RT      Augereau P., Mitalles F., Cavailles V., Gaudelot C., Parker M.,
RT      Rochereau H.;
RT      "Characterization of the proximal estrogen-responsive element of
RT      human cathepsin D gene.";
RL      Mol. Endocrinol. 8:693-703(1994).
RN      (6)
RP      SEQUENCE OF 170-180.
RX      TISSUE=Liver;
RT      Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
RT      Pasquall C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
RT      Appel R.D., Hughes G.J.;
RT      Submitted (JUN-1992) to the SWISS-PROT data bank.
RN      (7)
RP      VARIANT VAL-58.
RX      MEDLINE=20179010; PubMed=10716266;
RT      Papasotiriopoulos A., Bagli M., Kurz A., Kornhuber J., Forstl H.,
RT      Maier W., Pauls J., Lautenschlager N., Heun R.;
RT      "A genetic variation of cathepsin D is a major risk factor for
RT      Alzheimer's disease.";
RL      Ann. Neurol. 47:399-403(2000).
RN      (8)
RP      X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX      TISSUE=Spleen;

```

```

RX      MEDLINE=93223670; PubMed=8467789;
RA      Metcalf P., Fusek M.;
RT      "Two crystal structures for cathepsin D: the lysosomal targeting
RT      signal and active site.";
RL      EMBO J. 12:1293-1302(1993).
RN      (9)
RP      X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX      TISSUE=Liver;
RX      MEDLINE=93342076; PubMed=8393577;
RA      Balawin E.T., Bhat T.N., Gulnik S., Hosur M.V., Sowder R.C. II,
RA      Cachau R.E., Collins J., Silva A.M., Erickson J.W.;
RT      "Crystal structures of native and inhibited forms of human cathepsin
RT      D: implications for lysosomal targeting and drug design.";
RL      Proc. Natl. Acad. Sci. U.S.A. 90:6796-6800(1993).
CC      -1- FUNCTION: CATHEPSIN D IS AN ACID PROTEASE ACTIVE IN INTRACELLULAR
CC      PROTEIN BREAKDOWN.
CC      -1- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC      that of pepsin A. Does not cleave the 4-Gln-1-His-5 bond in B
CC      chain of insulin.
CC      -1- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC      -1- SUBCELLULAR LOCATION: Lysosomal.
CC      -1- POLYMORPHISM: The Val-58 allele is significantly overrepresented
CC      in demented patients (11.8%) compared with nondemented controls
CC      (4.9%). Carriers of the Val-58 allele have a 3.1-fold increased
CC      risk for developing AD than noncarriers.
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC      EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M11233; AAB59529.1; -
DR      EMBL; X05344; CAA28955.1; -
DR      EMBL; M63138; AAA51922.1; -
DR      EMBL; M63134; AAA51922.1; JOINED.
DR      EMBL; M63135; AAA51922.1; JOINED.
DR      EMBL; M63136; AAA51922.1; JOINED.
DR      EMBL; M63137; AAA51922.1; JOINED.
DR      EMBL; L12980; AAA16314.1; -
DR      EMBL; S74689; AAD14156.1; -
DR      EMBL; S52557; AAD13868.1; -
DR      PIR; A25771; KKHUDD.
DR      PDB; 1LYA; 31-JAN-94.
DR      PDB; 1LYB; 31-JAN-94.
DR      MEROPS; A01.009; -
DR      SWISS-2DPAGE; P07339; HUMAN.
DR      Sienna-2DPAGE; P07339; -
DR      MIM; 116840; -
DR      InterPro; IPR001969; Asp_protease.
DR      InterPro; IPR001461; Pepsin.
DR      Pfam; PF00026; asp; 1.
DR      PRINTS; PR00792; PEPSIN.
DR      PROSITE; PS00141; ASP_PROTEASE; 2.
KW      Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Signal; Zymogen;
KW      Polymorphism; Alzheimer's disease; 3D-structure.
FT      SIGNAL 1 18
FT      PROPEP 19 64 ACTIVATION PEPTIDE.
FT      CHAIN 65 412 CATHEPSIN D.
FT      CHAIN 65 161 LIGHT CHAIN (PROBABLE).
FT      CHAIN 169 412 HEAVY CHAIN (PROBABLE).
FT      ACT_SITE 97 97
FT      ACT_SITE 295 295
FT      DISULFID 91 160
FT      DISULFID 110 117
FT      DISULFID 286 290
FT      DISULFID 329 366
FT      CARBOHYD 134 134
FT      CARBOHYD 263 263

```

N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)

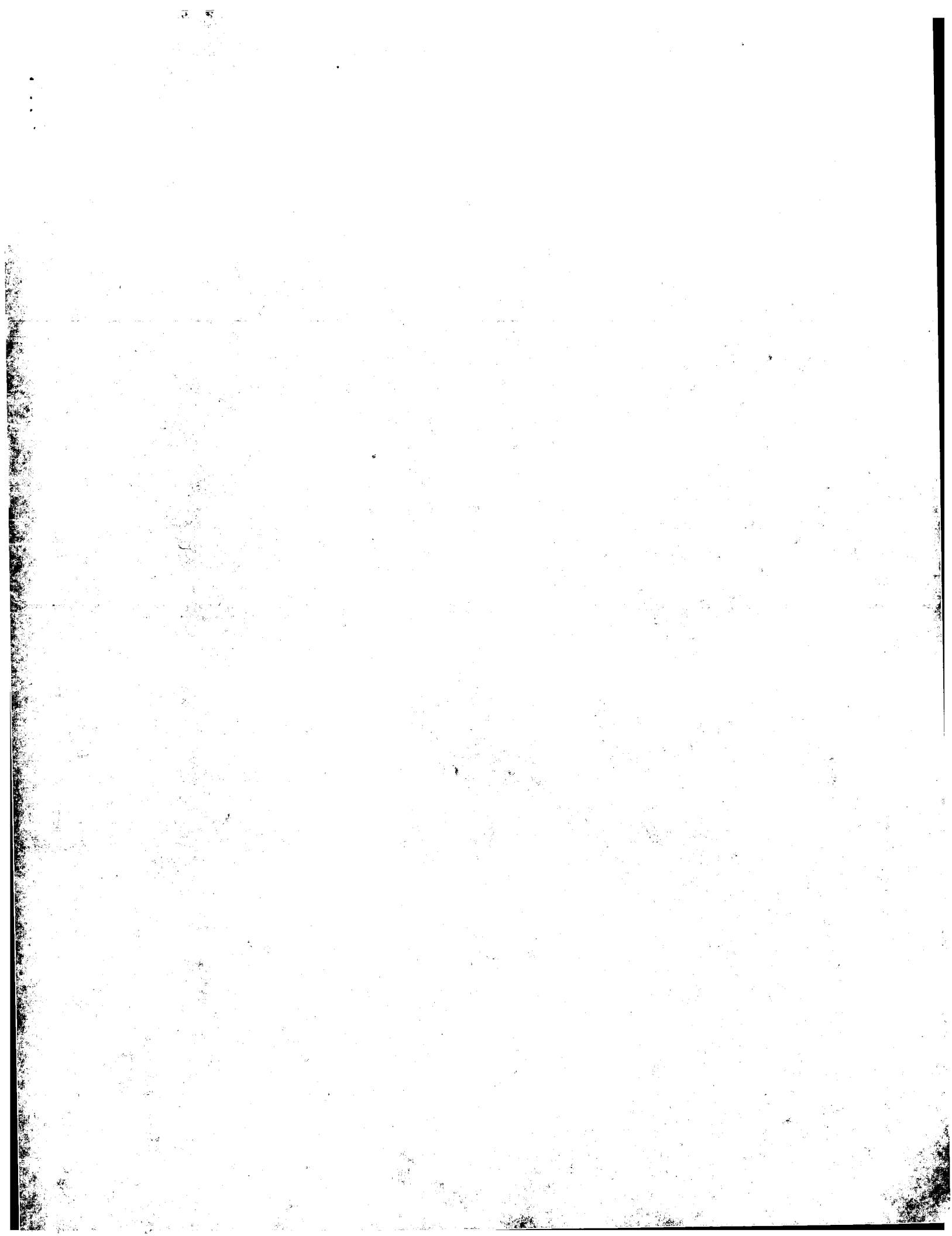
FT	VARIANT	58	58	A -> V (ASSOCIATED WITH INCREASED RISK IN AD; POSSIBLY INFLUENCES SECRETION AND INTRACELLULAR MATURATION; IN DBSNP:17571).
FT				/FTID=VAR_011621.
FT	STRAND	67	74	
FT	TURN	75	77	
FT	STRAND	78	85	
FT	TURN	86	89	
FT	STRAND	90	97	
FT	TURN	98	99	
FT	STRAND	103	107	
FT	TURN	108	109	
FT	STRAND	112	113	
FT	TURN	115	118	
FT	STRAND	119	123	
FT	TURN	123	123	
FT	STRAND	125	127	
FT	TURN	129	130	
FT	STRAND	132	141	
FT	TURN	146	158	
FT	STRAND	172	184	
FT	TURN	188	192	
FT	STRAND	197	200	
FT	TURN	204	206	
FT	STRAND	208	210	
FT	TURN	214	220	
FT	STRAND	221	222	
FT	TURN	223	233	
FT	STRAND	243	247	
FT	TURN	248	248	
FT	STRAND	252	254	
FT	TURN	255	263	
FT	STRAND	267	267	
FT	TURN	268	269	
FT	STRAND	270	279	
FT	TURN	280	281	
FT	STRAND	284	285	
FT	TURN	287	288	
FT	STRAND	290	294	
FT	TURN	296	297	
FT	STRAND	298	298	
FT	TURN	301	303	
FT	STRAND	305	315	
FT	TURN	316	316	
FT	STRAND	318	319	
FT	TURN	322	323	
FT	STRAND	325	328	
FT	TURN	329	334	
FT	STRAND	338	344	
FT	TURN	343	344	
FT	STRAND	345	349	
FT	TURN	351	354	
FT	STRAND	355	357	
FT	TURN	359	362	
FT	STRAND	365	368	
FT	TURN	370	372	
FT	STRAND	377	379	
FT	TURN	383	385	
FT	STRAND	387	390	
FT	TURN	391	392	
FT	STRAND	393	398	
FT	TURN	399	402	
FT	STRAND	403	409	

Query Match 13.7%; Score 295.5; DB 1; Length 412;
Best Local Similarity 28.5%; Pred. No. 4.2e-17;
Matches 105; Conservative 58; Mismatches 146; Indels 59; Gaps 15;

QY 30 YVEMTVGSPQTLNIIIVDQSSNFAGVAPHPFL-----HRYYORQLSSTYRDLRKGV 83
DB 79 YGEGIGTIPQCTTVVFDIGSSNLMVPSIHCKLLDIACWIIHKNYNDKSKSTYVKNKTSF 138

QY	84	YVPYTGKMEGELGDLVSP-----HGPNVTVRANIAATESDEKFFINGSNMEGI 134
DB	139	DIHGSSLSLGSYSDIVSPQOSASASALGSKYKEROFGGATNQPGITFLIAAFDCI 198
QY	135	LGLAVAEIARPPDLSLEPFPSLVKQTHV-PNLFSLDLCAGPPLNOSVLAISGSMITG 193
DB	199	LGMAYPRIS--VNNVLEVFNLMOQKLVDPNIFSFYL-----SRDDPAOGGGLMLG 248
QY	194	GIDHSLYTGLMTPTIREKYEYIIVREI-NGODLKMCKREYNKDKSIYDGGTINLRJ 252
DB	249	GTDSKYYKGSLSLVNTRKAWOVHLDQVEVASGLTL---CKE--GCEALVDGTISLMVG 303
QY	253	PKRYEAAVRSIKAASTKEFPDGFMLGEOLV-CWQAGTTPMNIIPVLSLYLMEVYTNOS 311
DB	304	PVDEVRELOKALGAVPLIQ-----GEYMIPOEKVST-----LPAITLKLGG---KG 346
QY	312	FRITILPQOYLRPEYDVATSDQCYKFAISO-----SSTGVAGAVIMEGFYVFPDRAR 366
DB	347	YKLS--PEDYTLKVSQAKTL--CLSGFMGMDIPPPSGPLMLGDVFIKRYTVYFDRDN 402
QY	367	RIGFAVSA 374
DB	403	RVGFAEAA 410

RESULT 15
PEP4_RABIT STANDARD; PRT; 387 AA.
ID PEP4_RABIT
AC P28713;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pepsin II-4 precursor (EC 3.4.23.1) (Pepsin A).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91009127; PubMed=2129536;
KA Kageyama T., Tanabe K., Koitani O.;
RT "Structure and development of rabbit pepsinogens. Stage-specific
RT zymogens, nucleotide sequences of cDNAs, molecular evolution, and
RT gene expression during development."
RL J. Biol. Chem. 265:17031-17038(1990).
CC -!- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS
CC INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC ALSO CLEAVED TO SOME EXTENT.
CC -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE PHE-, LEU-.
CC -!- DEVELOPMENTAL STAGE: PEPINOGENS IN GROUP I, II, AND III WHERE
CC THE PREDOMINANT ZYMOGENS AT LATE POSTNATAL STAGE
CC -!- MISCELLANEOUS: THE EXPRESSION OF PEPINOGEN GENES IS REGULATED BY
CC HORMONES AND RELATED SUBSTANCES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
DR PIR: D38302; D38302.
DR HSP: P00790; IFSN.
DR -MEROPS: A01.001; -.
DR InterPro: IPR001969; Asp-protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF000026; asp. 1.
DR PRINTS: PR00792; Pepsin.
DR PROSITE: PS00141; ASP_PROTEASE; 2
KW Hydroxylase; Aspartyl protease; Digestion; Stomach; Gastric juice;
KW zymogen; signal; Phosphorylation; Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 59 ACTIVATION PEPTIDE.
FT CHAIN 60 387 PEPsin II-4.
FT MOD_RES 129 129 PHOSPHORYLATION (POTENTIAL).
FT ACT_SITE 93 93 BY SIMILARITY.
FT ACT_SITE 276 276 BY SIMILARITY.
FT DISULFID 106 111 BY SIMILARITY.
FT DISULFID 267 271 BY SIMILARITY.



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2002, 09:17:12 ; Search time 50.29 Seconds

(without alignments)
1400.058 Million cell updates/sec

Title: US-09-724-571-58

Perfect score: 2156
Sequence: 1 ETDEEPEEPEGRGSEVEMVD.....GPFVTLDMEDCGYNIPQIDE 407

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP viirus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2156	100.0	532	4	Q9ULS1
2	2001.5	92.8	476	4	Q9BYC1
3	1898	88.0	457	4	Q9BYC0
4	1753.5	81.3	432	4	Q9BYC9
5	1149	53.3	266	11	Q9CUC5
6	1121	52.0	439	4	Q9H2V8
7	1116.5	51.8	514	11	Q9NZL1
8	967.5	44.9	396	4	Q9NZL2
9	933	43.3	468	4	Q9NZL2
10	712.5	33.0	213	4	Q9P0D2
11	566.5	26.3	253	11	Q9R1P7
12	332.5	15.4	391	5	Q9VAP6
13	332	15.4	354	5	Q9GYX7
14	312.5	14.5	386	6	Q9BGU5
15	308	14.3	384	13	Q9DEC2
16	308	14.3	385	13	Q9DEC4

17	305	14.1	386	6	Q9GMY7	Q9gmy7 rhinolophus
18	305	14.1	387	6	Q9GMY8	Q9gmy8 sorex unguis
19	304.5	14.1	372	5	Q9YIK3	Q9yik3 drosophila
20	304.5	14.1	387	13	Q9DDV5	Q9ddv5 salvelinus
21	304	14.1	387	6	Q9GMY9	Q9gmy9 suncus murel
22	303.5	14.1	383	13	Q9DE45	Q9de45 salvelinus
23	303	14.1	383	13	Q9DEC3	Q9dec3 xenopus lae
24	302.5	14.0	376	13	Q9PUR8	Q9pur8 pseudopleur
25	301	14.0	382	13	Q9PRG9	Q9prg9 gallus gall
26	301	14.0	423	5	Q9VKP7	Q9vkp7 drosophila
27	298.5	13.8	384	13	Q9I3J2	Q9i3j2 rana catesb
28	295.5	13.7	386	6	Q9GMY6	Q9gmy6 canis faml
29	294	13.6	396	13	Q934J8	Q934j8 chionodraco
30	290.5	13.5	381	6	Q9GK11	Q9gk11 camelus dro
31	288	13.4	399	13	Q934J8	Q934j8 podarcis si
32	287.5	13.3	444	5	Q21966	Q21966 caenorhabdi
33	284	13.2	398	13	P87370	P87370 oncorhynch
34	284	13.2	427	5	P91802	P91802 schistosoma
35	281	13.0	378	13	Q9PUR9	Q9pur9 pseudopleur
36	279.5	13.0	390	6	Q9GK10	Q9gk10 camelus dro
37	278	12.9	370	6	Q9TTW1	Q9ttw1 bos taurus
38	278	12.9	399	13	Q9DDB9	Q9ddb9 brachydantio
39	278	12.9	422	5	Q96906	Q96906 onchocerca
40	277.5	12.9	480	6	Q28950	Q28950 sus scrofa
41	277	12.8	446	5	Q9N9H3	Q9n9h3 neocator ame
42	276	12.8	389	13	Q9PWK1	Q9pwk1 gallus gall
43	276	12.8	396	13	Q9W643	Q9w643 gallus gall
44	276	12.8	396	13	Q9DEX3	Q9dex3 clupea hare
45	275	12.8	385	6	Q29080	Q29080 sus scrofa

ALIGNMENTS

RESULT	ID	Q9ULS1	PRELIMINARY;	PRT;	532 AA.
1	Q9ULS1	01-MAY-2000 (TREMBL)	13, Created)		
AC	Q9ULS1	01-OCT-2001 (TREMBL)	18, Last sequence update)		
DT	01-DEC-2001 (TREMBL)	19, Last annotation update)			
DE	KIAA1149 PROTEIN (FRAGMENT).				
GN	KIAA1149.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=BRIN;				
RX	MEDLINE=20039618; PubMed=10574461;				
RA	Hirosewa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;				
RT	*Characterization of cDNA clones selected by the Genemark analysis				
RT	from size-fractionated cDNA libraries from human brain.;				
RL	DNA Res. 6:329-336(1999).				
DR	EMBL; AB032975; BAA8463.2; -				
DR	HSSP; P56272; IAMS.				
DR	MEROPS; A01.004; -				
DR	InterPro; IPR001969; ASP-protease.				
DR	Pfam; PF00026; ASP; 1.				
DR	PRINTS; PR00792; PEPsin.				
DR	PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.				
FT	NON_TER				
SO	SEQUENCE	532 AA;	58720 MW;	988135D05FBD2E8 CRC64;	

Query Match 100.0%; Score 2156; DB 4; Length 532;
Best Local Similarity 100.0%; Pred. No. 3.3e-175;
Matches 407; Conservative 0; Mismatches 0; Gaps 0;
Indels 0; Gaps 0;
1 ETDEEPEEPEGRGSEVEMVDNLKSGCGYIVEMVGSPPOTLNLVDTGGSNFAVGAP 60
|||||

```

Db 77 ETDEPEEPGRGSEFVEMVDNLKSGGGYVEMTVGSPQTLNLTIVDTGSSNFAVGAAP 136
QY 61 HPFLHRYRQOLSTYRDLRKGVVPTTQCKMEGELGTDLVSIIPHGNVTVRANIAATE 120
Db 137 HPFLHRYRQOLSTYRDLRKGVVPTTQCKMEGELGTDLVSIIPHGNVTVRANIAATE 136
QY 121 SDFPFINSNMEGILGLAVAEIARPDLSLEPFDSLVKQTHVPLNLSIQCGAGFPLNOS 180
Db 197 SDFPFINSNMEGILGLAVAEIARPDLSLEPFDSLVKQTHVPLNLSIQCGAGFPLNOS 256
QY 181 EYLASVGGSMITGGIDHSLYTGSLMTYPIRREMYEVIIVAEINGODLKMDCKEYNDK 240
Db 257 EYLASVGGSMITGGIDHSLYTGSLMTYPIRREMYEVIIVAEINGODLKMDCKEYNDK 316
QY 241 SYVDSGTTNLRPKKVEFAVAKSIRKASTKEKPDGFMGLGEOLYCWMQAGTTPMNIFFPIS 300
Db 317 SYVDSGTTNLRPKKVEFAVAKSIRKASTKEKPDGFMGLGEOLYCWMQAGTTPMNIFFPIS 376
QY 301 LYLMGEVTVNOSFRITLLPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 360
Db 377 LYLMGEVTVNOSFRITLLPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 436
QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMDCGYNIPQIDE 407
Db 437 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMDCGYNIPQIDE 483

RESULT 2
Q9BYC1 PRELIMINARY: PRT: 476 AA.
ID AC Q9BYC1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BETA-SITE APP CLEAVING ENZYME I-476.
GN BACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA MEDLINE=21408467; PubMed=11516562;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production."
RL Neurosci. Lett. 307:9-12(2001).
DR EMBL; AB050436; BAB40931.1; -.
DR HSSP; P32329; IYPS.
DR InterPro; IPR001969; ASP_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 476 AA; 52907 MM; 6C8C87F8A953AF66 CRC64;

```

```

Query Match 92.8%; Score 2001.5; DR 4; Length 476;
Best Local Similarity 93.9%; Pred. No. 4,1e-162; Indels 25; Gaps 1;
Matches 382; Conservative 0; Mismatches 0;

```

```

QY 1 ETDEPEEPGRGSEFVEMVDNLKSGGGYVEMTVGSPQTLNLTIVDTGSSNFAVGAAP 60
Db 46 ETDEPEEPGRGSEFVEMVDNLKSGGGYVEMTVGSPQTLNLTIVDTGSSNFAVGAAP 105
QY 61 HPFLHRYRQOLSTYRDLRKGVVPTTQCKMEGELGTDLVSIIPHGNVTVRANIAATE 120
Db 106 HPFLHRYRQOLSTYRDLRKGVVPTTQCKMEGELGTDLVSIIPHGNVTVRANIAATE 165
QY 121 SDFPFINSNMEGILGLAVAEIARPDLSLEPFDSLVKQTHVPLNLSIQCGAGFPLNOS 180

```

```

Db 166 SDFPFINSNMEGILGLAVAEIARPDLSLEPFDSLVKQTHVPLNLSIQCGAGFPLNOS 200
QY 181 EYLASVGGSMITGGIDHSLYTGSLMTYPIRREMYEVIIVAEINGODLKMDCKEYNDK 240
Db 201 EYLASVGGSMITGGIDHSLYTGSLMTYPIRREMYEVIIVAEINGODLKMDCKEYNDK 260
QY 241 SYVDSGTTNLRPKKVEFAVAKSIRKASTKEKPDGFMGLGEOLYCWMQAGTTPMNIFFPIS 300
Db 261 SYVDSGTTNLRPKKVEFAVAKSIRKASTKEKPDGFMGLGEOLYCWMQAGTTPMNIFFPIS 320
QY 301 LYLMGEVTVNOSFRITLLPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 360
Db 321 LYLMGEVTVNOSFRITLLPQOYLRPVEDVATSDDCYKFAISOSTGTVMGAVIMEGFYV 380
QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMDCGYNIPQIDE 407
Db 381 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMDCGYNIPQIDE 427

```

```

RESULT 3
Q9BYC0 PRELIMINARY: PRT: 457 AA.
ID AC Q9BYC0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BETA-SITE APP CLEAVING ENZYME I-457 (BETA-SITE APP CLEAVING ENZYME
DE TYPE C).
GN BACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA MEDLINE=21408467; PubMed=11516562;
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production."
RL Neurosci. Lett. 307:9-12(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-EXOCRINE PANCREAS;
RA Zaccarelli D., De Pietri Tonelli D., Schnurbus R.;
RT "New beta-site APP cleaving enzyme isoform (BACE-1C) obtained from
RT human pancreas."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB050437; BAB40932.1; -.
DR EMBL; AF338817; AAK38375.1; -.
DR HSSP; P32329; IYPS.
DR InterPro; IPR001969; ASP_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 457 AA; 51068 MM; C794C9A9B85FE7A2 CRC64;

```

```

Query Match 88.0%; Score 1898; DR 4; Length 457;
Best Local Similarity 89.2%; Pred. No. 2,5e-153; Indels 44; Gaps 1;
Matches 363; Conservative 0; Mismatches 0;

```

```

QY 1 ETDEPEEPGRGSEFVEMVDNLKSGGGYVEMTVGSPQTLNLTIVDTGSSNFAVGAAP 60
Db 46 ETDEPEEPGRGSEFVEMVDNLKSGGGYVEMTVGSPQTLNLTIVDTGSSNFAVGAAP 105
QY 61 HPFLHRYRQOLSTYRDLRKGVVPTTQCKMEGELGTDLVSIIPHGNVTVRANIAATE 120
Db 106 HPFLHRYRQOLSTYRDLRKGVVPTTQCKMEGELGTDLVSIIPHGNVTVRANIAATE 145
QY 121 SDFPFINSNMEGILGLAVAEIARPDLSLEPFDSLVKQTHVPLNLSIQCGAGFPLNOS 180

```



```

Db 146 -----PDDSLPEFDSLVKQTHVNPFLSLQCGAGFPNLNS 181
QY 181 EVLASVGSMTIGIDHSLSLTGSLMTPIRREMYEVIIVRVEINODLKMCKEYNYDK 240
    |||||||
Db 182 EVLASVGSMTIGIDHSLSLTGSLMTPIRREMYEVIIVRVEINODLKMCKEYNYDK 241
QY 241 SIVDSGTTNLRPKKVEAAVKSIIKAASTKEKPPDGFMLGEOLVCMQAGTTPMNIFFVIS 300
    |||||||
Db 242 SIVDSGTTNLRPKKVEAAVKSIIKAASTKEKPPDGFMLGEOLVCMQAGTTPMNIFFVIS 301
QY 301 LYLMGEVTNQSFRITILLPOQYLRPVEDVATSODDCYKFAISOSSTGTGAVIMGEFYV 360
    |||||||
Db 302 LYLMGEVTNQSFRITILLPOQYLRPVEDVATSODDCYKFAISOSSTGTGAVIMGEFYV 361
QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTE 407
    |||||||
Db 362 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTE 408

```

```

RESULT 4
Q9BYB9 PRELIMINARY; PRT; 432 AA.
AC Q9BYB9:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE BETA-SITE APP CLEAVING ENZYME I-432.
GN BACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID:9606;
RN [1]
RC TISSUE=FROM N.A.
RC TISSUE=BRIN;
RX MEDLINE=21408467; PubMed=11516562;
R Tanahashi H., Tablira T.,
RT "Three novel alternatively spliced isoforms of the human beta-site APP
RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
RT production."
RL Neurosci. Lett. 307:9-12(2001).
RL EMBL; AB050438; BAB40933.1; -.
DR HSSP; P32329; IYPS.
DR InterPro; IPR001969; Asp_protease.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 432 AA; 48212 MW; 96FC81E6F0ED01B CRC64;

```

```

Query Match 81.3%; Score 1753.5; DB 4; Length 432;
Best Local Similarity 83.0%; Pred. No. 4.8e-141;
Matches 338; Conservative 0; Mismatches 0; Indels 69; Gaps 1;
QY 1 ETDEPEEPGRSGVVEVDNLGRKSGGYYVEMTVGSPPTLILIVDTGSSNFAVGAP 60
    |||||||
Db 46 ETDEPEEPGRSGVVEVDNLGRKSGGYYVEMTVGSPPTLILIVDTGSSNFAVGAP 105
QY 61 HPFLHRYQROLSTYRDLRKGVVVPYTGKMEGELGTDLSIFHGPNVYRANIAITE 120
    |||||||
Db 106 HPFLHRYQROLSTYRDLRKGVVVPYTGKMEGELGTDLSIFHGPNVYRANIAITE 145
QY 121 SDRFFINSMNEGILGLAYEAIARPDDSLVFKQTHVNPFLSLQCGAGFPNLNS 180
    |||||||
Db 146 -----LCGAGFPNLNS 156
QY 181 EVLASVGSMTIGIDHSLSLTGSLMTPIRREMYEVIIVRVEINODLKMCKEYNYDK 240
    |||||||
Db 157 EVLASVGSMTIGIDHSLSLTGSLMTPIRREMYEVIIVRVEINODLKMCKEYNYDK 246
QY 241 SIVDSGTTNLRPKKVEAAVKSIIKAASTKEKPPDGFMLGEOLVCMQAGTTPMNIFFVIS 300

```

```

Db 217 SIVDSGTTNLRPKKVEAAVKSIIKAASTKEKPPDGFMLGEOLVCMQAGTTPMNIFFVIS 276
QY 301 LYLMGEVTNQSFRITILLPOQYLRPVEDVATSODDCYKFAISOSSTGTGAVIMGEFYV 360
    |||||||
Db 277 LYLMGEVTNQSFRITILLPOQYLRPVEDVATSODDCYKFAISOSSTGTGAVIMGEFYV 366
QY 361 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTE 407
    |||||||
Db 337 FDRARRRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTE 383

```

```

RESULT 5
Q9CUC5 PRELIMINARY; PRT; 266 AA.
AC Q9CUC5:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE BETA-SITE APP CLEAVING ENZYME (FRAGMENT).
GN BACE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=BRIN;
RX MEDLINE=21085660; PubMed=11217851;
R Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
R Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
R Saito T., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
R Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
R Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
R Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
R Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
R Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
R Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
R Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
R Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
R Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
R Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
R Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
R Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,
R Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
R Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK014390; BAB29317.1; -.
DR MEROPS; A01.004; -.
DR MGD; MGI:1346542; Bace.
DR InterPro; IPR001461; Pepsin.
DR Pfam; PF00026; asp. 2.
FT NON_TER 1
SQ SEQUENCE 266 AA; 30201 MW; B913FDA8ADAB4238 CRC64;

```

```

Query Match 53.3%; Score 1149; DB 11; Length 266;
Best Local Similarity 99.1%; Pred. No. 7.7e-90;
Matches 215; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 191 IIGIDHSLSLTGSLMTPIRREMYEVIIVRVEINODLKMCKEYNYDKSIVDSGTTNL 250
    |||||||
Db 1 IIGIDHSLSLTGSLMTPIRREMYEVIIVRVEINODLKMCKEYNYDKSIVDSGTTNL 60
QY 251 RLKRVFEAAVKSIIKAASTKEKPPDGFMLGEOLVCMQAGTTPMNIFFVISLYLMGEVTNO 310
    |||||||
Db 61 RLKRVFEAAVKSIIKAASTKEKPPDGFMLGEOLVCMQAGTTPMNIFFVISLYLMGEVTNO 120
QY 311 SFRITILLPOQYLRPVEDVATSODDCYKFAISOSSTGTGAVIMGEFYVDFDRARRIGF 370
    |||||||
Db 121 SFRITILLPOQYLRPVEDVATSODDCYKFAISOSSTGTGAVIMGEFYVDFDRARRIGF 180

```

QY	371	AVASCHVHDEFRTAAVEGPEFTVLMDEDCGYINPOTDE	407
Db	181	AVASCHVHDEFRTAAVEGPEFTVLMDEDCGYINPOTDE	217
RESULT	6		
Q9H2V8			
Q9H2V8	PRELIMINARY:	PRT:	439 AA.
AC	Q9H2V8;		
DC	01-MAR-2001 (TReMBLrel, 16, Created)		
DT	01-MAR-2001 (TReMBLrel, 16, Last sequence update)		
DT	01-DEC-2001 (TReMBLrel, 19, Last annotation update)		
DE	CDA13.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=PHEOCHROMOCYTOMA;		
RA	Li Y., Huang Q., Peng, Y. Song H., Yu Y., Xu S., Ren S., Chen Z.,		
RL	Han Z.;		
RL	Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; AF212252; AAC41783.1; ..		
DR	HSSP; P00797; ZREN.		
DR	InterPro; IPRO01969; Asp.protease.		
DR	InterPro; IPRO01461; Pepsin.		
DR	Pfam; PF00026; asp. 3.		
DR	PRINTS; PR00792; PEPsin.		
DR	PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.		
SQ	SEQUENCE 439 AA: 46275 MM: 02E0DE05F11602 CRC64;		

Query Match	52.0%;	Score 1121;	DB 4;	Length 439;
Best Local Similarity	54.8%;	Pred. No. 3.9e-87;		
Matches 207;	Conservative 63;	Mismatches 104;	Indels 4;	Gaps 2;

[illegible]

RESULT	7	
Q9JL18		
ID	Q9JL18	PRELIMINARY;
AC	Q9JL18;	PRT, 514 AA.
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	

```

DT      01-DEC-2001 (TREMBLREL. 19, Last annotation update)
DE      ASPARTYL PROTEASE 1.
GN      BACE2.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Choi D.K., Sugano S., Sakaki Y.,
RT      "Molecular characterization of the mouse Aspl gene, a homolog of the
RT      human Aspl (Down Syndrome Region aspartyl protease).",
RL      Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF216310; AAF36599.1; -.
DR      HSSP: P32393; 1YPS.
DR      MEROPS: A01.041; -.
DR      MGD: MGI:1860440; Bace2.
DR      InterPro: IPR001969; Asp-protease.
DR      InterPro: IPR001461; Pepsin.
DR      Pfam: PF00026; asp; 3.
DR      PRINTS: PR00792; PEPsin.
DR      PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_2.
SQ      Protease.
SQ      SEQUENCE      514 AA;  55799 MW;  A70725F2C1D5B47 CRC64;

```

Query Match	51.8%;	Score 1116.5;	DB 11;	Length 514;
Best Local Similarity	-53.7%;	Pred. No. 1.2e-86;		
Matches 209;	Conservative 66;	Mismatches 109;	Indels 5;	Gaps 3;

[illegible]

OS Homo sapiens (Human) .
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase.";
RL Cytogenet. Cell Genet. 89:177-184(2000).
DR EMBL: AF168277; AAF3536.1; -.
DR HSP: P00797; ZREN.
DR MEROPS: A01.041; -.
DR InterPro: IPR001969; Asp.protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF00026; asp. 2.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 396 AA; 43013 MW; 5023A7AF391CEAC9 CRC64;

Query Match	44.98;	Score 967.5;	DB 4;	Length 396;
Best Local Similarity	54.68;	Pred. No. 4e-74;		
Matches 183;	Conservative 54;	Mismatches 85;	Indels 13;	Gaps 4;

[illegible]

RESULT	9	09NZL2	PRELIMINARY;	PRT;	468 AA.
09NZL2	09NZL2	09NZL2			
AC	09NZL2	09NZL2			
1-DT	01-OCT-2000	(TREMBLrel. 15, Created)			
1-DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)			
1-DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
1-DT	ASPAARYL	PROTEASE.			
1-DT	BACE2.				
1-DT	Homo sapiens (Human).				
1-DT	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
1-DT	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
1-DT	NCBI_Taxid=9606;				
1-DT	[1]				
1-DT	SEQUENCE FROM N.A.				
1-DT	MEDLINE=20422477; PubMed=10965118;				
1-DT	Solans A., Estivill X., de La Luna S.;				
1-DT	"A new aspartyl protease on 21q22.3, BACE2, is highly similar to				
1-DT	Alzheimer's amyloid precursor protein beta-secretase.";				
1-DT	Cytogenet. Cell Genet. 89:177-184(2000).				
1-DT	EMBL: AF188276; AAF35835.1; .				
1-DT	HSSP: P00797; 2REN.				
1-DT	MEROPS: A01.041; .				
1-DT	InterPro: IPR001969; Asp-protease.				
1-DT	InterPro: IPR001461; Pepsin.				
1-DT	Pfam: PF00026; asp. 3.				

DR PRINTS: PR00792; PEP5IN.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 468 AA; 50324 MW; 717E0920126A0142 CRC64

Query Match	43.3%	Score 933;	DB 4;	Length 468;
Best Local Similarity	46.5%;	Pred. NO. 4.5e-71;		
Matches 180;	Conservative 59;	Mismatches 94;	Indels 54;	Gaps 3

[illegible]

RESULT 10	
Q9P0D2	
ID	Q9P0D2
AC	Q9P0D2;
DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	HSPC104 (FRAGMENT).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID	9606;
NCBI_TaxID	[1]
SEQUENCE FROM N.A.	
TISSUE=CORD BLOOD;	
Zhang Q.H., Ye M., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,	
Fan H.T., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;	
"Human partial CDS cloned from cd34+ stem cells."	
Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.	
EMBL; AF161367; AAF28927.1;	
Interim; IPRO01461; pepsin.	
Pfam; PF00026; asp; 1.	
NON TER	
1	
SEQUENCE	213 AA; 24338 MW; EC9DD3FA31CFA835C CRC64;

Query Match	33.0 %	Score 712.5	DB 4	Length 213
Best Local Similarity	83.58	Pred. No. 87e53		
Matches 137	Conservative 4	Mismatches 12	Indels 11	Gaps 1
193	GGIDHSLYTSSWVTPRRRRMYEVVIVVEINSGDLKMDCKEYNYDKSIYDSGTTNRL	252		
1	GGIDHSLYTSSWVTPRRRRMYEVVIVVEINSGDLKMDCKEYNYDKSIYDSGTTNRL	60		

RESULT	12		
Q9VKP6			
ID	Q9VKP6	PRELIMINARY;	PTT; 391 AA.
AC	Q9VKP6;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	CG17134 PROTEIN.		
GN	CG17134.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		

[illegible]


```

RESULT 15
09DEC2 PRELIMINARY; PRT; 384 AA.
AC 09DEC2;
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE PEPSINOGEN A.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Ikuzawa M., Inokuchi T., Kobayashi K., Yasumasa S.;
RT "Molecular Cloning of Pepsinogens in Adult Xenopus laevis and Bullfrog
RT Rana catesbeiana."
RT Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
DR EMBL: AB045380; BAB20798.1; -.
DR HSSP: P00790; 1PSN.
DR MEROPS: A01.001; -.
DR InterPro: IPR001969; Asp.protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF00026; asp_1.
DR PRINTS: PR00792; PEPSTN.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
KW Aspartyl protease; Hydrolase.
SQ
SEQUENCE 384 AA; 41888 MW; 729DD2E7953D9072 CRC64;

Query Match 14.3%; Score 308; DB 13; Length 384;
Best local Similarity 27.5%; Pred. No. 6,1e-18;
Matches 99; Conservative 60; Mismatches 133; Indels 68; Gaps 16;

QY 30 YVEMTVGSPPTLNIIVDTGSSNFAVGAAPHPL-----HRYQROLSTYRDLRG 82
DB 72 YGTLISIGTPQETFEYFDGSAHLW---PSYCSQACSNHNRFPQOSTFOATNP 128
QY 83 YVPTGKMGELGTLVSIHPGNTVRAITAESDK-FEINGSNMGILGLAYAE 141
DB 129 VSIQGTGSMGFLGYDPLQY--GNIQISNQMGISESEPGSLYXSPDGLGLAFPS 185
QY 142 IARPDSDLPEPDSLVKQTHVP-NLFSIOLGAGFPLNOSVILASVGSMTIGIDHSLY 200
DB 186 IA--SSQATPFVDNMWGGGLPQNLFSYLSDDG-----QTGSYVLEFGVDNSY 233
QY 201 TGSILWPTIRREWEYEVIIIVEINGDPL--KMDCKEYNKSIYDSTNLRLPKKVE 258
DB 234 SGLNMWPLTAETWQITLDSVSIQVIAQSQC-----QALVDIGTSLMTGPSTPI 286
QY 259 AAVKSTKAASTERKPPDGLGEOLV-CWQAGTTPWNIPIVYISLYLMGEVTNOSFRITL 317
DB 287 ANIQNYIGASQDSN-----GQYINCNNISNMPTIVE-----TIN 321
QY 318 POQY-LRPVEDVATSDDDCYK-FAISQSSTGT---VMGAVIMEGFYVFPDRARRIGFA 371
DB 322 GVOYPLSPSAVYRONOGCGSSGFOAMNLPNTNSGDLMLIGDVFIRQYFTVFDRAANNVAYA 381

```

Search completed: August 7, 2002, 09:17:13
 Job time: 217 sec